THE JOHN CREERAR LIBRARY
MAY 1953

Journal of Experimental Psychology

ARTHUR W. MELTON, Editor
HUMAN RESOURCES RESEARCH CENTER
LACKLAND AIR FORCE BASE
SAN ANTONIO, TEXAS

CONSULTING EDITORS

Judson S. Brown, *State University of Iowa*
Cletus J. Burke, *Indiana University*
Robert M. Gagné, *Human Resources Research Center*
Frank A. Geldard, *University of Virginia*
James J. Gibson, *Cornell University*
Clarence H. Graham, *Columbia University*
David A. Grant, *University of Wisconsin*
Lloyd G. Humphreys, *Human Resources Research Center*
Donald B. Lindsley, *University of California, Los Angeles*
Neal E. Miller, *Yale University*
Edwin B. Newman, *Harvard University*
Kenneth W. Spence, *State University of Iowa*
Benton J. Underwood, *Northwestern University*
Delos D. Wickens, *Ohio State University*

Lorraine Bouthilet, Managing Editor

CONTENTS

Variables Affecting Sensitivity of the Human Skin to Mechanical Vibration: C. E. SHERRICK, JR. 273
Cutaneous Interaction Resulting from Simultaneous Electrical and Mechanical Vibratory Stimulation: J. A. VERNON 283
A New Method for Scaling Pain: P. SWARTZ 288
An Investigation of the "Randomness" of Threshold Measurements: M. WERTHEIMER 294
Foveal Simultaneous Brightness Contrast As a Function of Inducing- and Test-Field Luminances: A. L. DIAMOND 304
Motor Performance As a Function of Click Reinforcement: B. REYNOLDS AND J. A. ADAMS 315
Acquisition and Extinction of an Instrumental Response Sequence in the Token-Reward Situation: W. W. LAMBERT, E. C. LAMBERT, AND P. D. WATSON 321
The Effects of Verbal and Nonverbal Responses in Mediating an Instrumental Act: W. E. JEFFREY 327
Reinforcement and "Hypothesis" in the Discrimination Behavior of Chimpanzees: H. W. NISSEN, B. LEVINSON, AND J. W. NICHOLS 334
Approach-Avoidance in an Ambivalent Object Discrimination Problem: R. THOMPSON 341
The Effects of Single and Compound Classes of Anagrams on Set Solutions: I. MALTZMAN AND L. MORRISETT, JR. 345
Effects of Task Instructions on Solution of Different Classes of Anagrams: I. MALTZMAN AND L. MORRISETT, JR. 351
Studies of Distributed Practice: XI. An Attempt to Resolve Conflicting Facts on Retention of Serial Nonsense Lists: B. J. UNDERWOOD 355
The Memory Effect of Visual Perception of Three-Dimensional Form: H. WALLACH, D. N. O'CONNELL, AND U. NEISSER 360

American Psychological Association

Vol. 45 No. 5 May 1953

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

The JOURNAL OF EXPERIMENTAL PSYCHOLOGY is published monthly, two volumes per year, by the American Psychological Association, Inc. The subscription rate per volume is $7.50, or $15.00 annually. Single copies are $1.50. Subscriptions, orders, and business communications should be addressed to the American Psychological Association, Inc., 1333 Sixteenth St. N.W., Washington 6, D. C.

This JOURNAL publishes original experimental investigations which are intended to contribute toward the development of psychology as an experimental science. Studies with normal human subjects are favored over studies involving abnormal or animal subjects. Studies in applied experimental psychology or engineering psychology may be accepted if they have broad implications for experimental and theoretical psychology. Normally, articles of a length exceeding 15 printed pages cannot be accepted. However, an integrated series of studies accomplished simultaneously (e.g., most doctoral dissertations) must be presented in a single article, rather than in a series of articles.

Articles are published in the order of their receipt, except in rare circumstances. Authors who request them are supplied with 50 free offprints without covers. All of the cost of an author's alterations in galley proof is charged to the author, and one-half of the cost of tables, cuts, and formulas. Priority in publication is given to articles whose authors assume the full cost of publication.

Address all articles submitted for regular or priority publication to the editor, Arthur W. Melton, Human Resources Research Center, Lackland Air Force Base, San Antonio, Texas.

Manuscripts submitted to the JOURNAL *must* adhere to the conventions concerning reference citations, preparation of tables and figures, manuscript format, etc. as described in the Publication Manual of the American Psychological Association (*Psychol. Bull.*, 1952, 49 (4, Suppl.), 389–449. In particular, the organization of the manuscript should be text, references, footnotes, tables (one to a page), figure titles (typed on separate page), and figures. All typed material *must* be double-spaced with ample margins. The term "subject," with its various forms, is abbreviated *S*, *Ss*, *S*'s, and *Ss*'; analogous abbreviations are used for "experimenter" (*E*) and "observer" (*O*). When in doubt about the practices of the JOURNAL, authors should examine a recent issue.

Entered as second-class matter, February 6, 1937, at the post office at Lancaster, Pa., under the Act of March 3, 1879.

Acceptance for mailing at the special rate of postage provided for in paragraph (d-2), Section 34.40, P. L. & R. of 1948, authorized September 11, 1947.

Journal of
Experimental Psychology

VOL. 45, NO. 5 MAY, 1953

VARIABLES AFFECTING SENSITIVITY OF THE HUMAN SKIN TO MECHANICAL VIBRATION[1]

CARL E. SHERRICK, JR.

University of Virginia

The problem of the sensitivity of the human skin to mechanical vibration has stimulated research by experimental psychologists, clinical neurologists, and physicists for many years. In this, as in many problem areas in sensory psychology, more reliable techniques have been developed through the use of electronic devices, and more recent engineering research and development have made available to research workers electromechanical transducers exactly designed to job specifications (**16, 17**). Experimental analyses of sensitivity to mechanical vibration have centered around the usual problem areas in studies of the sense modalities. Investigations have been made of the absolute sensitivity to mechanical vibration, both as a function of frequency (**20, 9, 11, 12, 15**) and as a function of body area (**8, 9, 11**). In addition, studies of the Weber fraction for intensity (**15, 18**) and frequency (**4, 15, 19**) have been made, along with studies of adaptation (**1, 2, 14, 23**) and effects of temperature and anesthesia on sensitivity (**3, 24, 25**). The majority of these and many other studies of vibratory sensitivity are discussed in reviews of the literature by Geldard (**7, 10**).

The purpose of the present paper is to consider the role of the skin and underlying tissues as determiners of the form of the sensitivity curve for vibration as a function of frequency. Several investigators, among them von Békésy (**20**), Gilmer (**11**), Hugony (**12**), and Knudsen (**15**), have found that the sensitivity of the fingertip to vibration (stated as the minimum perceptible amplitude of motion) describes a U-shaped curve when plotted as a function of frequency. The minimum value for this function is in the neighborhood of .5 micron in the frequency region of 100–300 cy./sec. Figure 1 is a plot of this function taken from the data of the investigators mentioned above. It should be noted that von Békésy employed a stimulus which moved laterally across the skin,

[1] Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Division of Psychology, Department of Graduate Studies, University of Virginia. The deepest appreciation of the writer is here expressed to Dr. Frank A. Geldard for his able advice and direction in the course of this research.

This work was supported by a contract between the Office of Naval Research and the University of Virginia.

while all other investigators stimulated the skin by means of a contactor moving perpendicularly to the skin surface. Preliminary research by the writer has shown no significant difference between thresholds measured under the two conditions, however.

The form of the function appears to be fairly stable from day to day and *S* to *S*, and variations of body area and contactor sizes larger than 1-mm. diameter do not greatly affect it. Two interesting effects have been noted, however, in the use of very small contactors. Gilmer (11) found that stimulation of the fingertip with a needle .28 mm. in diameter caused the region of maximum sensitivity to shift upward to about 900 cy./sec. Geldard (9), in studying the frequency response of individual pressure-sensitive "spots" on the volar forearm, found that thresholds were independent of frequency for the usual range studied (64–1024 cy./sec.). A consideration of the form of the sensitivity curve suggests that the skin and underlying tissues may behave as a simple mechanical system under forced vibration, i.e., that at some frequency of vibration the system attains a maximum velocity under a constant force. This frequency is the natural period of the system, and is in addition the point at which maximum energy is transferred to the tissues, other conditions being equal. If this hypothesis is correct, then the increased sensitivity of the human *S* to vibration occurs as a result of the spatial summation of impulses caused by the increased spread of vibratory disturbances through the tissues. The two effects mentioned above are in line with this hypothesis, since the increased tension produced in Gilmer's experiment by the small contactor could alter the mechanical characteristics of the tissues so as to decrease the natural period of the system.

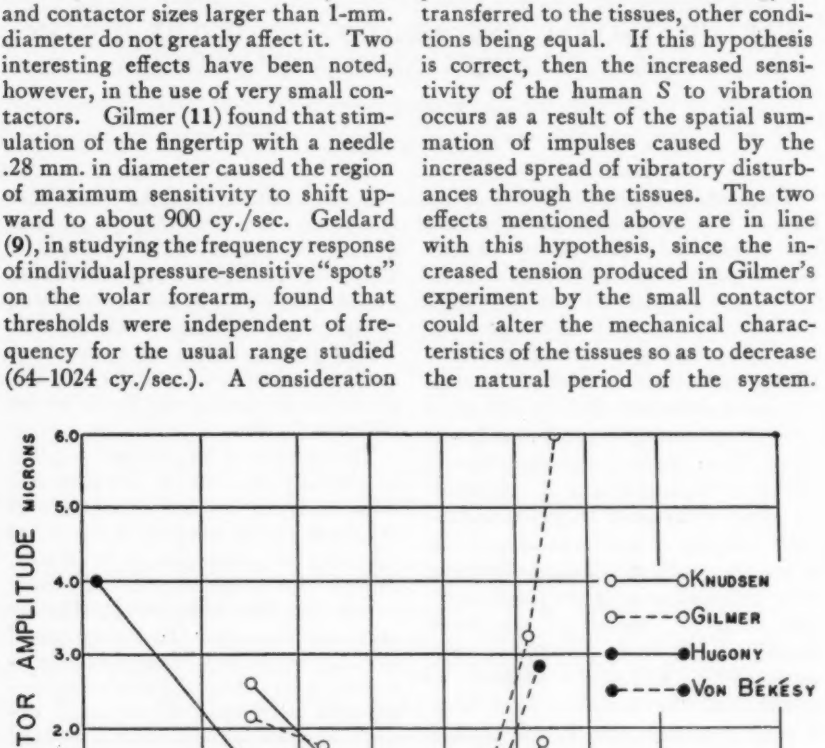

FIG. 1. Comparison of the results of four investigators for amplitude thresholds as a function of frequency for vibration. One micron = 10^{-6} meters.

The results of Geldard's experiment form one of the assumptions of the hypothesis, in that single pressure-sensitive "spots" or dense clusters of receptors do not show a natural period in the frequency range ordinarily studied. This fact is disputed by results obtained by von Békésy (**21**), who found in a similar study that vibration-sensitive spots exhibit the usual **U**-shaped sensitivity curve. Von Békésy stimulated the spots near hair stumps, however, while Geldard intentionally avoided such regions (**9**, p. 275). Jackson (**13**), who attempted to repeat and extend Geldard's work, found that spots appeared to show the usual dependence on frequency, but there is some doubt as to the adequacy of the technique used in determining the locus of the pressure-sensitive spots.

One additional fact would bear mention at this point. It is possible that the skin itself does not resonate mechanically in accordance with the hypothesis, but that instead bony tissue underlying the skin acts to reflect the waves of vibratory disturbance. If the bony tissue had a natural period in the frequency range of 100–300 cy./sec., the disturbance of the tissue system would be maximized, and sensitivity heightened in this range. An experiment by von Békésy (**22**) appears to strengthen this hypothesis. Measurements of the amount of damping of vibratory disturbances by six cervical vertebrae taken from a human cadaver showed that the maximum conductivity of the system fell near 250 cy./sec. An easily testable deduction from this hypothesis is that tissue relatively free of bone should not show the same curve form in its frequency function as would tissue overlying bone. A preliminary experiment was performed to make this test, and to obtain representative sensitivity curves for the apparatus and techniques used in the present research.

Experiment I

Apparatus.—The apparatus was similar in arrangement and elements to that used by nearly all previous investigators. A General Radio, Type 613-B, battery operated, beat-frequency oscillator drove a Thordarson T-31W50 amplifier, the output of which was led to an electromagnetic speaker driver unit which was similar in construction to the RCA Model 100-A driver unit. The power delivered to the driver unit was varied by means of a Selectohm, Model 501, potentiometer, and voltage across the coil of the driver unit was measured by a Ballantine, Model 300, vacuum tube voltmeter. Calibration of amplitude of motion of the contactor attached to the driver-unit armature was made by observing the motion of the contactor microscopically, aided by the flickering illumination of a General Radio, Model 631-A, Strobotac. This procedure permits accurate calibration and observation of wave form and direction of motion of the contactor (**8**). The contactor fastened to the armature extension of the driver unit was a brass knob 3.7 mm. in diameter.

Procedure.—Four adult male *O*s were tested for sensitivity of the right index fingertip using the apparatus described. After four to five training sessions of an hour each, five readings at frequencies of 40, 100, 200, 250, 300, 400, 500, 600, 800, and 1200 cy./sec. were obtained by a modification of the method of limits, using only the ascending series to avoid adaptation effects. The fleshy pad of the right index fingertip of *O* was stimulated. As a test of the hypothesis concerning the relative sensitivities of bony and bone-free tissues, the same procedure was carried out on the tip of *O*'s tongue. Five *O*s were tested, four of whom had been tested in the previous experiment.

Results.—Median values are plotted in Fig. 2 in comparison with the sensitivity function for the fingertip previously obtained. This comparison does not reveal the relative forms of the curves, however, and Fig. 3 better illustrates this relationship. The ordinate for each curve is expressed as the common logarithm of the ratio of the threshold values to the threshold at 200 cy./sec.

Comparison of the curves in Fig. 3 demonstrates that bony tissue cannot

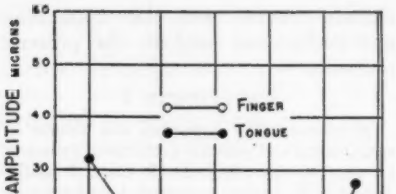

FIG. 2. Plot of absolute thresholds for five *Os* and two areas—tongue and right index fingertip—against frequency of vibration

be the only factor contributing to the minimum in the sensitivity curve, since the tongue has a curve form quite similar to that of the fingertip. Two possible hypotheses might be advanced to account for this fact: (*a*) The skin alone possesses mechanical characteristics which could produce resonance in the region of 200 cy./sec.; (*b*) the receptors themselves possess a sensitivity which is selective with respect to frequency, as von Békésy (**21**) and Jackson (**13**) found.

EXPERIMENT II

Purpose.—The evidence presented in the preceding section must be considered equivocal as facts in support of the hypothesis of tissue resonance. An alternative method of attack on this problem would be direct analysis of the spread of vibration from the origin of disturbance to points distant from the origin. The hypothesis of tissue resonance would predict that the spread of disturbance should be greater at frequencies in the range of maximal sensitivity. Moreover, by selecting areas of the body differing considerably in the amount of tissue separating skin surface and bone, some indication may be obtained of the differential conductivities of skin-and-soft-tissue and skin-and-bony-tissue systems.

Apparatus.—Serious difficulties were encountered in using as vibration pickups the piezo-electric and magnetic reluctance pickups designed for phonograph reproduction, since the frequency range of greatest importance to the experiment is a region of poor response for these devices. An alternate method is that of using the human *O* as a "pickup" by having him place his fingertip on the area to be tested. This procedure was tested on models which simulated the skin, and quite consistent and accurate results were obtained.

Procedure.—The areas of the body studied were: (*a*) the right volar forearm extending from a point 6 cm. distal from the olecranon of the ulna, along the midline of the forearm; (*b*) the region overlying the crest of the left tibia extending distally from a point 4 cm. distal from the apex of the patella. The vibratory stimulus was the contactor previously used, driven by the equipment described in the first experiment. One *S* served merely as a site for stimulation, while the *O* was employed as the vibratory "pickup." The areas on the *S* described above were selected and marked with semipermanent ink, and a series of four marks 1.0 cm. apart were placed on a line running distally from the original

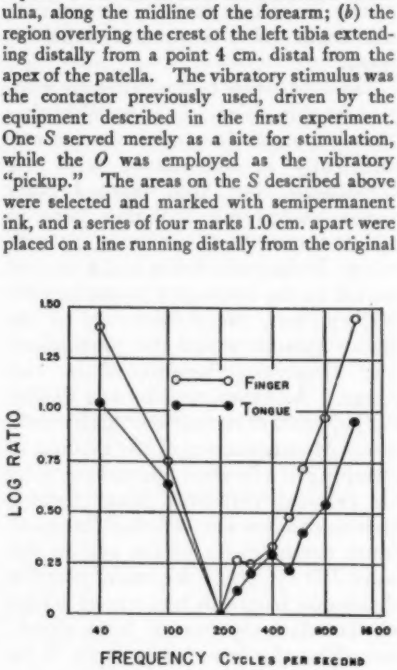

FIG. 3. Plot of relative thresholds for five *Os* and two areas—tongue and index fingertip—against frequency of vibration to threshold value at 200 cy./sec.

mark. The marks on the arm were kept on the median line, and those on the leg on the crest of the tibia. These same points could be re-established quite accurately at subsequent experimental sessions. The procedure for the experiment proper was to place the contactor on the origin point of *S*'s arm or leg. The distance to be tested and the frequency to be used were selected by *E* at random, and *O* placed his finger on the designated point. The *E* and *O* then followed the procedure previously used in determining thresholds. This procedure was repeated for five trials in ascending series for each of five frequencies, 100, 200, 250, 300, and 600 cy./sec., and for each of the four distances from the origin of disturbance. The *O*'s thresholds were also determined in the usual manner either at the beginning or the end of the main experimental session.

Results.—Curves are plotted in Fig. 4, showing the relative amplitude of disturbance at two distances, 1.0 and 4.0 cm., for the two areas as a function of frequency. The ordinate is the common logarithm of the ratio of *O*'s thresholds at the points on the *S*'s skin to the *O*'s thresholds on the contactor. The dip in the curves at 200 cy./sec. indicates that conductivity of the skin is heightened at this frequency. Moreover, the bony area appears to conduct vibration better at nearly all frequencies. This direct analysis of the variation of vibratory propagation on the skin with frequency lends more support to the resonance hypothesis, and the differential conductivity of the two areas studied provides additional evidence that the two tissue systems may produce the sensitivity curves obtained by past investigators.

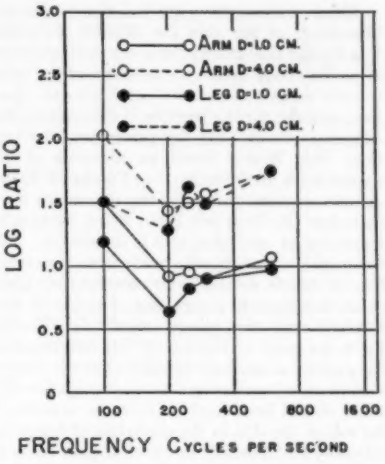

FIG. 4. Plot of conductivity of skin of volar forearm and of skin of leg overlying the tibia. Ordinate is $\log_{10}$ ratio of *O*'s threshold when acting as a "pickup" to threshold on contactor.

EXPERIMENT III

Purpose.—A consideration of the differential conductivity of the skin for various frequencies of vibration and of the effect of anatomical structure on vibrotactile sensitivity has demonstrated a possible qualitative relationship between physical characteristics of the skin and the sensitivity curve normally obtained in vibratory research. There exists a single physical characteristic which may conceivably encompass both these concepts, and which may be quantitatively described by methods at present available. This construct is mechanical impedance, a complex quantity derived from the characteristics of the system in question by the proper mathematical combination of the mass, elasticity, and friction of the system. If the skin may be represented by such a system, it follows that its impedance will depend on the frequency of forced vibrations acting on it. Moreover, at that frequency of the system representing its natural period the impedance will be at a minimum value. From the hypotheses stated previously, we should expect the impedance of the skin to exhibit a minimum value at the same point as do the amplitude threshold curves previously discussed.

Other workers have studied the mechanical impedance of the skin for different purposes. Von Békésy (20) attempted to use the impedance of various body areas as a criterion of the most suitable area for measures of sensitivity to vibration, and also for the purpose of determining the adequate stimulus for the perception of vibration. Von Békésy found no resonance of the tissues in his investigation, but Franke (5) found definite resonant points for the skin which were quite low (20–38 cy/sec.), but which increase in frequency as contactor area is decreased. The latter measurements were made over the rectus femoris muscle, an area so far removed from bony tissue that rigorous comparison of data with von Békésy's and the present results is difficult. From the work of Franke and his collaborators has evolved a thorough description of the skin as a medium of mechanical wave propagation (6) which should be of further aid in the analysis of the role of the skin in the mediation of pressure, vibratory and possibly mechanical pain stimuli.

Apparatus.—A method for obtaining the mechanical impedance of the skin was suggested in a paper by Wiggins (26). This author described the construction of and theory behind a mechanical impedance bridge which could be used to obtain the impedance of such devices as phonograph pickups, speaker driver units, diaphragms, and other apparatus used in sound production. The application of this device to the measurement of skin impedance required no elaborate changes in the bridge as originally designed, and rapid and accurate measures of skin impedance are easily obtainable.

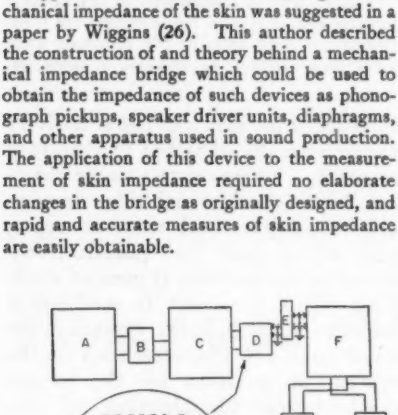

FIG. 5. Block diagram of apparatus for measuring skin impedance. Inset shows schematic view of reed-driver unit system.

The theoretical considerations related to the design of this apparatus may be stated simply: a vibrating driver unit drives at a known frequency two identical mechanical systems the impedances of which are known. Each system is constructed in such a manner that the force exerted on it by the driver unit produces a voltage output which may be amplified and read by means of a voltmeter. The two voltage outputs are made to cancel one another, and one of the systems is loaded with the unknown impedance. The differential effects on the two systems as given by the voltages from them, combined with functions of their known impedances, will give directly the value of the unknown impedance. This value may be obtained at any number of frequencies within the limitations of the device.

A diagram of the apparatus is shown in Fig. 5. The inset in this figure shows the reed-driver-unit system used in measuring impedance. The General Radio oscillator and Thordarson amplifier described in previous sections were used as a source of alternating current to excite an RCA 100-A driver unit which drove the reed system. A condenser voltage source and preamplifier, a two-channel amplifier, a vacuum tube voltmeter, and an oscillograph were employed to record the voltage outputs of the reeds. The reed was constructed of aluminum and, for low frequency measurements, had dimensions of 15.2 cm. length by 1.1 cm. width by .32 cm. thickness. The reed was clamped tightly at its center to a polystyrene support from which it was separated by .01 in. by means of a brass shim. The clamping of the reed produced two reeds whose ends were free to vibrate, and whose new dimensions were $5.7 \times 1.1 \times 0.32$ cm. each. The surfaces of the polystyrene facing the ends of the reeds were coated with Aquadag to make them electrically conductive, and screws fastened in the polystyrene led from these surfaces to wires going to the preamplifier. The reeds, shielding surfaces, etc. were all connected to electrical ground of the system. Each end of the reed-polystyrene unit then represented a condenser, one plate of which moved with respect to the other plate, in this case the Aquadag-coated polystyrene representing the stationary plate relative to the reed. The driver-unit contactor shaft drove the polystyrene support and the reeds, and the design of the polystyrene support was such that it vibrated only as a unit. The reeds, however, being clamped at one end and having their mass distributed uniformly along their length, vibrated continuously in and out of phase with the polystyrene support. A high direct current voltage was placed across the reed and Aquadag-coated polystyrene, and the vibrations of the system then produced a voltage in the same manner as a condenser microphone. Wires lead-

ing from the reed and the polystyrene were sent to the condenser input preamplifier and thence to the two-channel amplifier. In this manner the alternating voltages from the two reeds were amplified separately, and a switching arrangement permitted *E* to read the voltage output and wave form of either reed singly or both together. A phase-shifting device inserted at the output of the two-channel amplifier permitted combination of the reed voltages in any desirable phase, e.g., to cancel the signals completely or to add them maximally. The cathode ray oscillograph was employed to insure that only sinusoidal voltages were used.

Procedure.—The procedure for determining skin impedance was to set the contactor in motion at an amplitude near threshold for the frequency of vibration used. The signals from the reeds were adjusted at equal voltage values by manipulating amplifier controls, and the phase of the signals shifted so that they cancelled one another completely. The fingertip or knuckle was then placed on the edge of one reed as shown in Fig. 5 and the voltage change recorded. This voltage represented the difference in voltage between loaded and unloaded reeds. A second reading of the voltage on the unloaded reed alone completed the measurement. The ratio of the difference in voltage between loaded and unloaded reeds to the voltage of the unloaded reed alone was then multiplied by a constant obtained from calculations of the mass, compliance, and frequency of vibration of the reed. This operation gave a value which was the mechanical impedance of the tissue for an effective area of .35 sq. cm., the area of the edge of the reed used. This procedure was repeated over a frequency range of 90–1000 cy./sec., at frequencies which produced reed output voltages sinusoidal in character. The latter precaution prevented measurements at multiples of the input frequency, which might give incorrect values for skin impedance. Because of such distortion, frequencies below 90 cy./sec. could not be employed.

Results.—The results of this experiment are shown in Fig. 6. One *S* was employed who had a normal sensitivity curve for vibration, the right index fingertip and the knuckle of the second joint of the right index finger being taken as areas for study. Each point on the curves represents a median of three readings. The curves have a minimum value at 150–200 cy./sec., rising from this point to higher values on either side of the minimum. The form of these curves leaves little doubt that the skin and underlying tissues resonate in the frequency region of greatest sensitivity to vibration. In addition it appears that the bony tissue represented by the knuckle has a sharper resonance point than the fingertip. Readings of the impedance of the tongue were taken, but the response of this area below 300 cy./sec. was such that the recording instruments would not give reliable measures. It appears from the readings obtained that the tongue possesses a low impedance at all frequencies, rising slowly at the high frequencies to values near 1,000 mechanical ohms. If the tongue does resonate, the curve is very broadly tuned as compared with the fingertip or the knuckle.

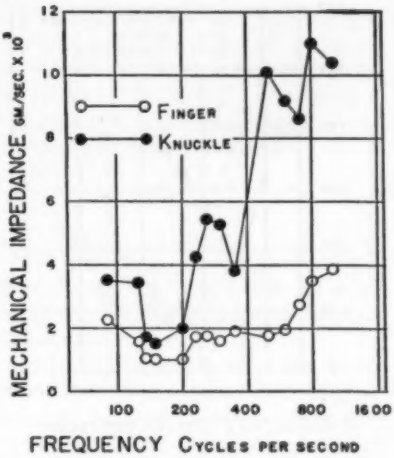

FIG. 6. Plot of mechanical impedance, in gm./sec., as a function of frequency of vibration for an area of .35 sq. cm. on the skin of the right index fingertip and the knuckle of the second joint of the right index finger

A comparison of the curves of sensitivity of vibration and mechanical impedance of the fingertip is made in

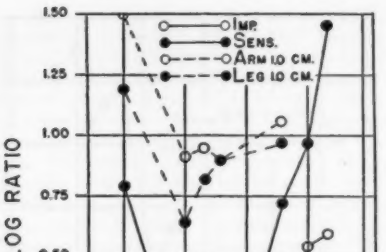

FIG. 7. Comparison of skin impedance and sensitivity of the finger, and skin conductivity of the arm and leg, as a function of frequency of vibration. Ordinate for impedance is $\log_{10}$ ratio of impedance to impedance at 200 cy./sec. Ordinate for sensitivity is same as Fig. 3, for conductivity same as Fig. 4.

Fig. 7. The curve for mechanical impedance is plotted as the common logarithm of the ratio of impedance to impedance at 200 cy./sec., and the curve for sensitivity is the same as Fig. 3. While deviations between these curves are apparent at the lower and higher frequencies, it is notable that the curves follow one another quite closely in the middle range. For purposes of additional comparison the "skin" and "bone" conductivity are plotted in the same manner as in Fig. 4. The deviation of the skin impedance curve from the sensitivity curve may be explained by an examination of the relationship between impedance, force, and velocity in the vibrating system. Impedance is equal to the ratio of the force acting on the vibrating system to the velocity of the system. The force acting is equal to the product of the mass, amplitude, and square of the angular velocity of the system (angular velocity = $2\pi \times$ frequency of vibration). The velocity of the system is equal to the product of its amplitude and angular velocity. The amplitude threshold curve has been shown to be a nonlinear function of frequency, and frequency is, of course, an independent variable. It is obvious from this brief mathematical analysis that impedance is not a linear function of amplitude of vibration, and hence the two curves could not be expected to coincide at all points. The important point to be noted in the comparison of the two curves is the existence of minima at the same frequency, and the common locus of points of inflection at 300 cy./sec. The point of inflection may indicate a second resonant point in the system, which lends additional support to the hypothesis proposed earlier, viz., that there are actually two mechanical systems, skin and bony tissue, and that these systems have resonant points at different frequencies. A more careful analysis of the response of radically different types of tissue should determine the validity of this hypothesis.

DISCUSSION

The findings of the present research appear to support the hypothesis that the selective sensitivity of the human *O* to vibratory disturbances is a result of the mechanical characteristics of the skin and underlying tissues, assuming that receptors for the vibratory stimulus do not themselves respond selectively to various frequencies of vibration, as Geldard (**9**) found. These characteristics appear to be such that the tissues resonate at a frequency of about 200 cy./sec., producing a spatial summation of receptor impulses and effectively increasing the sensitivity to vibration. Certain results of previous investigators ap-

pear to lend some credence to the present hypothesis. For example, Gilmer (11) found that stretching the skin of the fingertip by means of mechanical devices or by use of small contactors producing steep gradients served to shift the point of maximum sensitivity toward higher frequencies, as was previously stated. In addition to this, the work of Hugony (12) appears to provide some support. Hugony found "equal loudness" contours for vibrations of various frequencies and intensities at the fingertip. The point of maximal sensitivity for each contour shifted from about 200 cy./sec. near the absolute threshold to about 100 cy./sec. at high intensity levels (300 times threshold at 200 cy./sec.). The findings of Franke (5) indicate that increasing the area of disturbance on the skin resulted in a fall of the resonance point of the tissue to lower frequencies, presumably because mass of the resonant system increased. The increase in intensity in Hugony's experiment would serve to involve greater masses of skin also, owing to the conductivity of the tissue. It would then be predicted from the hypothesis presently advanced that the resonance point, and hence the point of maximal sensitivity, would fall to lower frequencies.

It would be unrealistic to avoid mention of experiments which do not agree with results noted in preceding paragraphs. It has been stated that neither von Békésy (21) nor Jackson (13) obtained results similar to Geldard's (9), and reasons have been offered for the difference in results. Von Békésy's failure to find resonance of the skin in his impedance measures appears as the strongest contradiction to the present results, but there is the possibility that von Békésy was measuring the impedance of extremely small areas of tissue. Such measures might well show no such variation in impedance as the present work, at least in the frequency range of interest in vibrotactile research. Only further research can reveal the causes of the contradictory results which are in the literature.

Summary

1. An examination was made of the role of the skin as a mechanical system in producing the usual U-shaped sensitivity curve to vibration as a function of frequency.

2. It was assumed that the receptors sensitive to mechanical vibration do not respond selectively to frequency and that either skin or bony tissue possesses mechanical characteristics such that it has a natural period of vibration, accompanied by maximum spread of disturbance, in the frequency region of 100 to 300 cy./sec.

3. Hypotheses based on this assumption were as follows:

(*a*) If skin tissue does not have a natural period but bony tissue does, the relative sensitivities of regions such as the tongue and fingertip should differ at various frequencies.

(*b*) The conductivity of the skin should be maximal at the frequency of greatest sensitivity to vibration, and bony tissue should conduct vibration more efficiently than relatively bone-free areas.

(*c*) The mechanical impedance of tissue should be minimal at the frequency of greatest sensitivity, and the impedance of bony tissue should be relatively different from that of tissue further removed from bone.

4. These hypotheses were tested experimentally, and it was concluded that:

(*a*) Both skin and bony tissue may have resonant points within the usual frequency range of vibrotactile investigation.

(*b*) The conductivity of skin is maximal at the frequency of greatest vibratory sensitivity, and bony tissue is more efficient in the propagation of disturbances.

(*c*) Mechanical impedance is minimal at the frequency of greatest sensitivity, and bony tissue shows a more sharply tuned resonance curve than relatively bone-free areas.

5. The results of other investigators were discussed in the light of the present research, and it was concluded that more extensive research on sensitivity of pressure "spots" and on the physical characteristics of the skin may reveal the cause of the lack of agreement of results at present in the literature.

(Received September 2, 1952)

References

1. Bellows, R. M. An experimental isolation of some factors determining response to rhythmic cutaneous stimulation: I. Frequency, pressure, and time. *J. exp. Psychol.*, 1936, **19**, 716–731.
2. Cohen, L. H., & Lindley, S. B. Studies in vibratory sensibility. *Amer. J. Psychol.*, 1938, **51**, 44–62.
3. Cummings, S. B. The effects of local anesthesia on tactile and vibratory thresholds. *J. exp. Psychol.*, 1938, **23**, 321–338.
4. Dunlap, K. Palmesthetic difference sensibility for rate. *Amer. J. Physiol.*, 1911, **29**, 108–114.
5. Franke, E. K. Mechanical impedance of the surface of the human body. *J. appl. Physiol.*, 1950–51, **3**, 582–590.
6. Franke, E. K., Gierke, H. E. von, Oestreicher, H. L., Parrack, H. O., & Wittern, W. W. von. Physics of vibrations in living tissues. Wright Air Development Center, Dayton, Ohio. A. F. Technical Report No. 6367, 1951.
7. Geldard, F. A. The perception of mechanical vibration: I. History of a controversy. *J. gen. Psychol.*, 1940, **22**, 243–269.
8. Geldard, F. A. The perception of mechanical vibration: II. The response of pressure receptors. *J. gen. Psychol.*, 1940, **22**, 271–280.
9. Geldard, F. A. The perception of mechanical vibration: III. The frequency function. *J. gen. Psychol.*, 1940, **22**, 281–289.
10. Geldard, F. A. The perception of mechanical vibration: IV. Is there a separate "vibratory sense"? *J. gen. Psychol.*, 1940, **22**, 291–308.
11. Gilmer, B. von H. The measurement of the sensitivity of the skin to mechanical vibration. *J. gen. Psychol.*, 1935, **13**, 42–61.
12. Hugony, A. Über die Empfindung von Schwingungen mittels des Tastsinnes. *Z. Biol.*, 1935, **96**, 548–553.
13. Jackson, M. M. The frequency function of vibratory sensitivity. Unpublished master's thesis, Univer. of Virginia, 1939.
14. Kampik, A. Experimentelle Untersuchungen über die praktische Leistungsfähigkeit der Vibrationsempfindungen. *Arch. f. Psychol.*, 1930, **76**, 3–70.
15. Knudsen, V. O. "Hearing" with the sense of touch. *J. gen. Psychol.*, 1928, **1**, 320–352.
16. Lewis, R. C. Applying shakers for product vibration analysis. *Product Eng.*, 1950, **21**, 150–154.
17. Lewis, R. C. Test techniques using electrodynamic shakers. *Product Eng.*, 1951, **22**, 142–147.
18. Périlhou, P., & Piéron, H. Quelques caractéristiques des sensations vibratoires. *C.R. Soc. Biol. Paris*, 1942, **136**, 448–449.
19. Roberts, W. H. A two-dimensional analysis of discrimination of differences in the frequency of vibrations by means of the sense of touch. *J. Franklin Inst.*, 1931, **213**, 286–312.
20. von Békésy, G. Über die Vibrationsempfindung. *Akust. Zeit.*, 1939, **4**, 316–334.
21. von Békésy, G. The neural terminations responding to stimulation of pressure and vibration. *J. exp. Psychol.*, 1940, **26**, 514–519.
22. von Békésy, G. Über die Stärke der Vibrationsempfindung und ihre objektive Messung. *Akust. Zeit.*, 1940, **5**, 113–124.
23. Wedell, C. H., & Cummings, S. B. Fatigue of the vibratory sense. *J. exp. Psychol.*, 1938, **22**, 429–438.
24. Weitz, J. Vibratory sensitivity as affected by local anesthesia. *J. exp. Psychol.*, 1939, **25**, 48–64.
25. Weitz, J. Vibratory sensitivity as a function of skin temperature. *J. exp. Psychol.*, 1941, **28**, 421–436.
26. Wiggins, A. M. Mechanical impedance bridge. *J. acoust. Soc. Amer.*, 1943, **15**, 50–53.

Journal of Experimental Psychology
Vol. 45, No. 5, 1953

CUTANEOUS INTERACTION RESULTING FROM SIMULTANEOUS ELECTRICAL AND MECHANICAL VIBRATORY STIMULATION[1]

JACK A. VERNON[2]

University of Virginia

In the history of human sensory investigations electrical currents have frequently been employed as stimulating agents. Unfortunately, however, the degree of understanding of the action of electrical stimuli on human tissue is by no means proportional to the extent to which the technique has been employed. One of the most persistent confusions lies in the inability of investigators to designate the locus of the electrical stimulation, that is to say, whether the electrical stimulus activates specialized sensory receptor cells or whether it by-passes these receptor mechanisms and produces direct excitation of the nerve.

Von Frey (**10**) performed punctate electrical stimulation of the skin which yielded the sensation that was characteristic for the spot stimulated. That is, when electrically aroused, warm spots responded with sensations of warmth, cold spots yielded coldness, etc. For von Frey this meant that the locus of electrodermal stimulation could be either the endorgan or its nerve fibers. Of these two possibilities, Gildemeister (**7**), favored the nerve fibers. In support of this position he claimed that stimulation by electrodes placed over the nerves leading to the skin produced the same sensation as when the electrodes were placed on the skin. He also pointed out that when the skin was anesthetized, pressure sensitivity was reduced, whereas electrical sensitivity remained unaltered. Fischer (**5**) has also found that anesthetizing the skin with ethyl chloride had no effect on electrical sensitivity.

The research to be reported here was specifically designed with the problem of the locus of cutaneous electrical stimulation in view. Preliminary exploration verified the earlier observations of von Frey (**10**) that threshold and slightly suprathreshold intensities of alternating current delivered to the fingertips produce the sensation of vibration. Moreover, these electrically aroused vibratory sensations are indistinguishable from those produced by mechanical vibratory stimulation. Thus it seems plausible that these two forms of energy excite the same peripheral mechanisms.

If the same mechanisms are involved, then a *simultaneous* application of electrical and mechanical vibratory forces to a *common locus* on the fingertips should produce an interaction. The form of this interaction will depend upon the nature of the peripheral processes and also on their phase relation. If the electrical stimulation produces a mechanical vibration of the skin and this mechanical vibration is present in the same sensory cells that are being subjected to a direct mechanical stimulation, then there will be a pure vectorial summa-

[1] Submitted in partial fulfillment of the requirements of the Ph.D. degree in the Division of Psychology, Department of Graduate Studies, University of Virginia. The writer is greatly indebted to Dr. Frank A. Geldard for his guidance and encouragement throughout this research. This work was done under Contract N7-onr-372 T.O. II, between the University of Virginia and the Office of Naval Research.

[2] Now at Princeton University.

tion of the two vibrations. The result of the summation will depend upon the phase relations between the two series of waves. If the waves are of the same frequency and intensity, then an in-phase relation (0° phase relation) will give a summation in which the effect will be double that of one of the stimuli alone. An out-of-phase relation (180° phase relation) will give a complete cancellation of the stimulation. Other phase relations will give intermediate forms of summation. These effects of summation should be obtained at any frequency to which the skin is capable of responding. It should also be possible to produce cutaneous beats by slightly mistuning the frequency of the two stimuli.

That electrodermal stimuli produce vibratory movements of the skin is claimed by Meyer (**9**), Kahler and Ruf (**8**), and Chocholle (**1, 2, 3, 4**). These investigators found that a tone, audible to observers outside the electrical circuit, was produced when electrical currents were passed through the skin. They further found that the necessary condition for the production of sound was a dry-type electrode placement. Chocholle (**1**), in particular, has demonstrated that this electrostatic property of the skin is not sufficient to produce tone when wet-type electrodes are utilized. Persistent attempts in this laboratory, explained in more detail later, have failed to support the finding of the above investigators.

If the electrical stimulus does not impart a vibratory movement to the skin but arouses sensation by sensory or neural stimulation, the interaction resulting from simultaneous mechanical and electrical vibratory stimuli delivered to a common locus on the skin may be something other than a simple vectorial summation. However, some form of summation or inhibition, or both, may be expected. It is difficult to predict the outcome of this more complex type of interaction because of the complicated and little-known time factors involved in endorgan reception, conversion, and transmission. Such an interaction, whatever its precise form, ought to result in intensity variations perceptible as beats when the frequency of the two stimuli is slightly mistuned.

The experiments to be reported here were designed to test the possibility of any form of interaction between electrical and mechanical vibratory stimulation.

Method

It was necessary to construct two separate stimulation systems, one to provide mechanical vibratory stimulation and another to provide stimulation of the skin directly by alternating currents. These two systems were electrically isolated so that the only common pathway of the two stimuli was within the skin of the observer.

The mechanical vibratory stimulus was supplied by an oscillator, amplified by a 50-w. amplifier the energy from which was converted into mechanical vibration by an RCA Type 100-A speaker motor. The mechanical vibratory stimulus was delivered to the skin by a plastic rod, 5.0 mm. in diameter and 2.2 cm. in length, which was attached to the armature of the speaker motor. A slightly concave brass cap, 5.1 mm. in diameter, was threaded into the tip of the plastic stylus. This brass cap served as the mechanical contactor and also as the "active" electrode for the electrical stimulus. The speaker motor was so pivoted as to insure a fairly constant pressure gradient of the skin of the fingertip on the stylus.

The amount of the mechanical vibratory stimulation was expressed as microns of excursion of the stylus. This extent of movement was determined by a stroboscopic optical system described by Geldard (**6**).

The electrical stimulus was provided by a separate oscillator and power amplifier and was applied to *O* at the active electrode. The circuit was completed by an "indifferent" electrode (a flat plastic cup 25 mm. in diameter) placed under the heel of the hand. The concavities of the brass cap and the plastic cup were filled with electrode paste to insure intimate electrical contact between the skin and the electrodes. The

intensity of the electrical stimulus was always expressed in terms of the product of the generated current and voltage. At all times both types of stimulation were of sinusoidal waveform.

A monitoring oscilloscope which received signals from each system allowed *E* to observe the established beat rate and thereby check the accuracy of *O*'s report. Inasmuch as the detection of cutaneous beats was difficult and *O* sometimes confused his pulse with the beats, it was necessary to require *O*s to detect and identify a particular beating rate. This rate was then shifted at some point in time without *O*'s knowledge, and he was required to detect the shift and identify the new beating rate. This procedure worked well and was used throughout the investigation of cutaneous beats.

Results

Four *O*s detected and correctly identified cutaneous beats under the conditions just described when the two stimuli were in the region of 250∼. The phenomenon was reported as a surging of intensity of the cutaneous pattern. At times the surging was clearly evident in this region of frequency, but at other times it was absent. In a given observation period beats would appear and disappear in a spasmodic fashion that corresponded to no obvious manipulation of the conditions. The clearness of the beats or their ease of detection was quite unpredictable.

In an effort to ferret out the optimal conditions for the phenomenon of cutaneous beats, there was manipulation of many variables such as intensive level, beating rate, order of presentation, loci on the fingertip, mechanical pressure gradient, and frequency of the stimuli. It was found that when the frequency of stimulation was within the range of about 280∼ to 310∼, when the intensities of the two types of vibratory stimuli were approximately equal and only slightly suprathreshold, and when the beating rate was from one to three beats per second, *clear* and *stable* cutaneous beats were produced. Of these three conditions the most critical requirement for consistent beats was that of the frequency range of stimulation.

The occurrence of these stable beats served as a qualitative indication of cutaneous interaction, but they did little to reveal the nature of the interaction. Once it was known that electrical and mechanical vibration would interact, it was possible to design a second and more quantitative experiment which would reveal whether the interaction was one of pure vectorial summation or one of a more complex nature. This second experiment was designed to measure the magnitude of the interaction. This was done by determining the effect of simultaneous *subliminal* electrical stimulation on mechanical vibratory thresholds under two conditions. In one condition the two stimuli, as measured at their source, were in phase, and in the other condition they were out of phase.

A subliminal electrical stimulus was employed so that the mechanical vibratory threshold could be detected, since the patterns resulting from supraliminal intensities of the two types of stimulation feel alike and would be confused with each other, thus rendering the desired measurement impossible.

To conduct this investigation it was necessary to modify the apparatus so as to present stationary in-phase and out-of-phase conditions rather than the continually changing phase condition utilized in the production of beats. This was accomplished by employing a single oscillator whose amplified output was divided across two matched transformers which led to the two stimulating channels. To produce and maintain the desired phase relationship between the electrical and mechanical stimuli, a double-pole, double-throw relay was installed in the mechanical stimulus circuit. When activated, this relay reversed the polarity between the two types of stimulation so that in one position the two stimuli were "in phase" and in the other position the two stimuli were "out of phase." It is necessary to re-emphasize that these phase relationships are those determined, by Lissajous' figures, between the sources of the two stimuli.

To avoid providing *O* with clues as to the

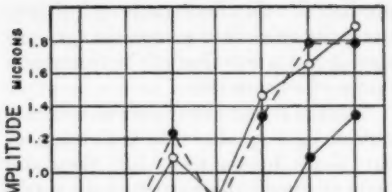

FIG. 1. Measurement of the magnitude of interaction between simultaneously presented electrical and mechanical vibratory stimuli. The open circles represent normal mechanical vibratory thresholds, the solid circles joined by a solid line respresent mechanical vibratory thresholds determined in the presence of an in-phase subliminal electrical stimulation, and the solid circles joined by a dashed line represent mechanical vibratory thresholds determined in the presence of an out-of-phase subliminal electrical stimulation.

nature of the stimulating conditions, the following procedure was adopted. Prior to each determination of a mechanical vibratory threshold the electrical stimulus was increased from zero to an intensity level permitting the first clear identification of vibration, and then it was gradually decreased to a subliminal value. Sometimes the subliminal value was maintained and a mechanical vibratory threshold was determined in its presence (i.e., a "combined" threshold), while at other times, randomly decided, the electrical stimulus was decreased to zero and a normal mechanical vibratory threshold was determined. Before utilizing this last step, it was first demonstrated that just previous electrical stimulation had no effect upon mechanical vibratory thresholds. Thus, as a result of the adopted procedure, *O* had no knowledge as to when subliminal electrical stimulation was present or what phase relation between the two stimuli had been established. The precaution was also taken to prevent cues based upon presentation time.

The above procedure was utilized with three sophisticated *O*s at stimulation frequencies of 250 ~ through 310 ~ at 10 ~ intervals which were presented in random order. The normal mechanical vibratory threshold was taken to be the mean value of 28 determinations at each frequency of stimulation. The combined thresholds were the means of 14 determinations at each frequency and for each of the two phase conditions. The intensity of the subliminal electrical background for the combined thresholds was about 70% of the normal electrical threshold as measured by volt-amperes. Frequent rest periods were introduced to prevent subject fatigue.

The results for one *O*, which are typical for all three *O*s, are presented in Fig. 1. From Fig. 1 it can be seen that when the two stimuli are "in phase," a form of summation occurs which effectively increases mechanical vibratory sensitivity. That this interaction is not a simple one is indicated by the lack of any type of summation when the stimuli are "out of phase."

As already pointed out, the actual phase relation between these two stimuli at the site of interaction is unknown. First, there is a time lag in the mechanical vibratory stimulus. This delay is produced by the speaker motor which converts electrical energy into mechanical vibratory movements (electro-mechanical transduction). Second, different temporal delays are introduced for the two stimuli as a function of their conduction rate through human tissue.

The transducer lag was measured and found not to be uniform over the frequency range of 250~ to 300~. At 300~ the lag was 135°, while at the other frequencies it was only 70°. Yet in spite of these lag variations, the cutaneous interaction illustrated in Fig. 1 is of a fairly uniform nature at the frequencies indicated. This type of uniformity would not have been expected for an interaction that was the result of a simple vectorial summation. These results, therefore, cast doubt on the claim that alternating currents stimulate the skin by first developing vibratory movements of the skin. This point was investigated further with an RCA Vibration Pick-

up, Type TMV-150B, coupled to a General Radio Amplifier, Type 814-A. This pickup system was placed on the fingertip 2 mm. distant from the active electrode (used dry, without electrode paste). The sensitivity of the pickup system depends upon the amplification factor coupled to it. In this case it was rendered sufficiently sensitive to detect a mechanical vibration intentionally delivered to the skin which was of so low an amplitude as to be always subliminal. Yet when electrical current alone was delivered to the skin, the pickup system did not indicate any mechanical movements on the skin even when unbearably high electrical currents were used.

Summary

The simultaneous application of electrical and mechanical vibratory forces to a common locus on the fingertip was demonstrated to produce a cutaneous interaction. This interaction, as indicated by the detection of cutaneous beats, was best produced in the narrow frequency range of about 280~ to 310~.

The magnitude of the cutaneous interaction was measured by determining the effect of simultaneous subliminal electrical stimuli upon mechanical vibratory thresholds. When the two forces were presented in one phase relation (the absolute value was unknown), the mechanical vibratory thresholds were greatly lowered as compared with normal mechanical vibratory thresholds, but when the two forces were presented in a phase relation 180° different from the first, there was little or no departure from normal thresholds.

The evidence shows that electrical stimuli do not stimulate the skin by first producing mechanical movements in the skin.

In that an interaction between electrical and mechanical vibratory stimuli has been demonstrated, even under limited conditions, it now appears fruitful to plan future work. The investigation is being extended by the use of continuous phase variation.

(Received September 2, 1952)

References

1. Chocholle, R. Effet de la polarisation sur la sensation auditive engendrée par stimulation de l'épiderme au moyen d'un courant alternatif. *C. R. Soc. Biol. Paris*, 1948, **142**, 1393–1395.
2. Chocholle, R. Emission d'ondes acoustiques audibles par l'épiderme sous l'action d'un courant électrique alternatif. *C. R. Soc. Biol. Paris*, 1948, **142**, 469–471.
3. Chocholle, R. Sensation auditive engendrée par stimulation électrique de l'épiderme céphalique au moyen d'un courant alternatif. *C. R. Soc. Biol. Paris*, 1948, **142**, 467–469.
4. Chocholle, R. Comparaison des variations des seuils de l'effet électrophonique en fonction de la fréquence dans diverses conditions. *C. R. Soc. Biol. Paris*, 1950, **144**, 36–38.
5. Fischer, H. Ueber den Einfluss der Hauttemperatur auf die Prickelempfindung. *Z. Sinnesphysiol.*, 1937, **67**, 149–158.
6. Geldard, F. A. The perception of mechanical vibration: II. The response of pressure receptors. *J. gen. Psychol.*, 1940, **22**, 271–280.
7. Gildemeister, M. Untersuchungen über die Wirkung der Mittelfrequenzströme auf den Menschen. *Pflüger's Arch. ges. Physiol.*, 1944, **247**, 366–404.
8. Kahler, H., & Ruf, F. Untersuchungen über das Hören durch Schwingungsanregung in Körper auf elektrostatischem Wege. *Z. Hals-Nas.-u. Ohrenheilk*, 1931, **29**, 218–230.
9. Meyer, M. Hearing without cochlea? *Science*, 1931, **73**, 236–237.
10. von Frey, M. Physiologische Versuche über das Vibrationsgefühl. *Z. Biol.*, 1915, **65**, 417–427.

Journal of Experimental Psychology
Vol. 45, No. 5, 1953

A NEW METHOD FOR SCALING PAIN[1]

PAUL SWARTZ[2]

University of Rochester

Prior to the late 1940's, scant attention had been paid to the quantitative aspects of pain intensity. For some years Wolff and his co-workers (1) had reported judgments of pain intensity in terms of 1+ to 10+, 1+ signifying a just noticeable pain and 10+ the most intense pain ever "experienced" by *S*. Libman (9) exerted pressure on the styloid process or squeezed the biceps muscles. Patients compared the induced pain with a "spontaneously" occurring pain. Estimates of pain intensity based on such procedures as these have not contributed very much towards a solution of the general problem, since they utilized stimuli incapable of precise quantification.

In 1947 Hardy and his collaborators (6) set about constructing a pain scale based upon the discrimination of differences in intensity of measurable pain stimuli. By means of the radiant-heat technique these investigators established the *dol* scale of pain intensity. Further work by these same investigators (7) resulted in a pain scale determined by a fractionation procedure which coincided with the original scale based on the summation of j.n.d.'s.

Lately, the practical usefulness of the dol scale was tested by Hardy and Javert (4). By a bracketing technique the pain intensity of childbirth was rated against the intensity of pain produced by exposing the person's skin to radiant heat.

Although the dol scale is a landmark in experimentation on pain, its general usefulness is limited by the restricted nature of the scaling technique used in determining the actual mathematical characteristics of the scale. Hardy (7) employed a passive rather than active fractionation technique. The *S*s merely reported the intensity of predetermined test stimuli as fractions of the initial standard stimulus. They did not actively determine by a trial and error process actual fractional values of a standard stimulus.

In the light of this procedural limitation the present investigation was designed to provide more rigorous information concerning the quantitative aspects of the pain response. The pain stimulus consisted in 60-cycle alternating current. The site of contact was the tooth pulp in a lower molar. The advantages of this algesimetric technique over the radiant-heat method have been well stated by Goetzl, Burrill, and Ivy (3).

METHOD

Subjects.—Forty-five undergraduates at the University of Rochester served as *S*s. Twenty-four were males, whose ages ranged from 18 to 25 yr. with a mean age of 20.2 yr. Twenty-one were females; their ages ranged from 18 to 34 yr. with a mean age of 20.1 yr.

Each of the 45 *S*s served in two experimental sessions. From three to seven days intervened between sittings.

Apparatus.—A circuit diagram of the apparatus employed is presented in Fig. 1. The two micro-safety switches, one in each line of the

[1] The present study is a portion of a thesis submitted to the Department of Psychology of the University of Rochester in partial fulfillment of the requirements for the Ph.D. degree. The writer is indebted to Dr. G. R. Wendt for his generous advice and help during the planning and administration of this investigation. Dr. Wendt designed the pain circuit of the apparatus.

[2] Now at Hobart and William Smith Colleges.

pain circuit, served to protect *S*. The microammeter had a range of 500 microamperes. Readings above this level were made possible by placing a shunt in the circuit.

The timing unit was adapted from the Hunter-Brown Electronic Interval Timer (8). This provided time intervals ranging from .2 to 8.0 sec. The actual stimulus duration used was 1 sec., varying less than 2% in either direction.

When the two micro-safety switches are depressed, the current and voltage are registered on their respective meters but no current passes through *S*. When the timing switch is thrown and the relay opens, current passes through *S* and the ammeter reading drops slightly owing to the added resistance of *S*. This reduced current is the effective current, the amount actually stimulating *S*. In the actual experiment the safety and timing switches were operated only by *S* and never by *E*.

The units comprising the pain circuit were housed in a plywood box. The timing unit was installed in a metal container fastened to the side of the pain unit. On the *S*'s side of the plywood box were: (*a*) the two safety switches, (*b*) the timing switch, (*c*) the voltage control knob, encircled by a scale with continuous markings to indicate 2-v. steps (no units appear on the scale to indicate to *S* the actual voltage), (*d*) a knob controlling the 100,000 ohm variable resistor, and (*e*) the two electrode leads. On *E*'s side of the equipment were: (*a*) the voltmeter and ammeter, (*b*) the shunt switch, (*c*) a knob controlling the variable resistor, and (*d*) a bank of switches controlling the three 100,000 ohm resistors.

Two electrodes were used in this study, one for the tooth and the other for the chin. The tooth electrode consisted in a thin silver wire spring mounted on the crossbar of a rubber-dam tooth clamp and insulated from it. A piece of rubber dam was punctured and stretched over the crossbar until it rested on the two prongs of the clamp. The clamp was placed on a lower molar containing a silver amalgam filling on the occlusal surface. The tip of the wire coil pressed firmly on the filling. The piece of rubber dam was then stretched and passed under the clamp, fitted on the tooth and stretched across the front of the mouth. It was held in place by a rubber-dam retainer. This is the standard technique used in dental practice for fitting rubber dams. The rubber dam served to isolate the tooth and keep it dry. The indifferent chin electrode was a curved brass plate. It was coated with elec-

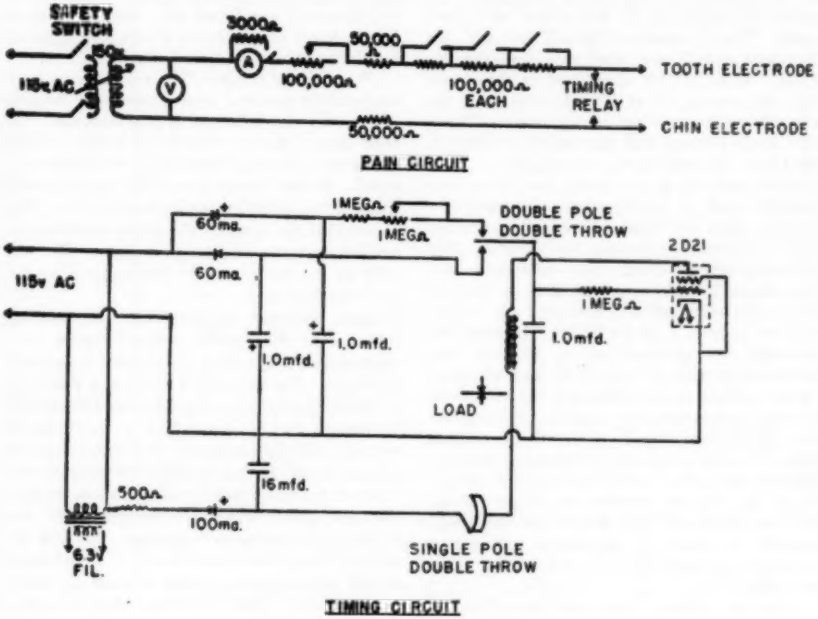

FIG. 1. Circuit diagram of pain stimulator

trode paste and fitted under the chin. It was held in place by an elastic band stretched over the top of the head. The tooth was dried by blasting air from a hand syringe. A saliva ejector was then placed in the mouth.

First session.—The *S* was placed in series with 100,000 ohms and the voltage set at zero. For this and subsequent readings the procedure was as follows: The *E* said "ready" and *S* depressed the safety switches. Then *E* said "now" and *S* threw the timing switch. After each stimulus presentation *S* was asked (*a*) whether he perceived the stimulus, (*b*) whether it was painful, and (*c*) whether he would submit to greater stimulus intensities. The *S* increased the voltage in 2-v. steps. Ten seconds intervened between stimulus presentations and each stimulus lasted 1 sec. Readings in microamperes were taken at the point of first perception of stimulation, first perception of pain, and at tolerance level, i.e., the maximum amount of pain to which *S* would voluntarily submit. To reduce the possibility of tissue damage no *S* was allowed to exceed a tolerance level of 1500 microamperes. The difference between the tolerance level and the pain threshold constituted the pain range.

Following the determination of the tolerance level a 2-min. rest period preceded the second half of the session. In this half *S* successively bisected his pain range by the method of adjustment. The *S* stimulated himself first with his threshold and 15 sec. later with his tolerance value. With 100,000 ohms fixed in the circuit and the current set at *S*'s tolerance limit the variable resistor was set at zero ohms. The *S* then made settings with the variable resistor in the 15-sec. interval between stimulations. Since the knob was set at zero ohms, the first setting could be made in one direction only, towards a stimulus intensity weaker than the tolerance limit. Unless *S*'s tolerance limit was at 1500 microamperes, he could make subsequent settings above his tolerance limit if he so desired. After each stimulation *S* corrected his setting until he judged it to be halfway between his threshold and tolerance values. He was then restimulated with his threshold and tolerance values to check on his judgment. If he wished, he could then make a new setting. If *S* reached the 100,000-ohm end of the variable resistor without having found the halfway point, another 100,000-ohm resistor was placed in the circuit by *E* and the variable resistor set at zero ohms. The two additional 100,000-ohm resistors could similarly be placed in the circuit, giving *S* a working range of 400,000 ohms without adjusting the variac.

After the halfway point was determined, *S* was stimulated with his threshold value again and then with the judged halfway value of the preceding pair of stimuli. The *S* then bisected this distance. In this manner *S* successively bisected his pain range until he either reported an inability to bisect further or set a judged halfway point: (*a*) at the upper limiting stimulus, (*b*) at the lower limiting stimulus, (*c*) above the upper limiting stimulus, or (*d*) below the lower limiting stimulus.

Second session.—In the second session the first half of the experimental session was the same as that in the first session. The same tooth was stimulated and a pain range determined.

The scaling procedure in this session differed in two important respects from that of the first session. First, the order of presentation of the limiting stimuli was reversed. The *S* was always stimulated first with his tolerance limit or previous halfway point, then with his threshold. Secondly, the bisection was always done with 500,000 ohms initially in the circuit, and with the voltage adjusted so that the current corresponded to *S*'s threshold.

In the actual scaling *S* increased the intensity of the current from his threshold towards the upper limiting stimulus. The bisections were made in the same direction as in the first fractionation session. First, the halfway point between tolerance and threshold was found. Then the halfway point between the first midpoint and the threshold was found, etc., until *S* could no longer bisect or made one of the four types of terminal settings mentioned above.

Treatment of results.—The results have been analyzed for men and women separately because of a possible sex difference in pain response. On both days following the determination of the pain-range threshold adaptation was frequently noted. It was therefore necessary to determine a new threshold before scaling could begin. This meant that for most *S*s the range bisected was smaller than the original pain range. We will refer to the range actually fractionated as the fractionating range.

Each bisection reduced the fractionating range. The successively reduced ranges were expressed as proportions of the total fractionating range. For example, if an *S* has a threshold of 100 microamperes and a tolerance level of 200 microamperes, this gives him a fractionating range of 100 microamperes. If he bisects at 160 microamperes, he has reduced his range to six-tenths of the fractionating range. Thus, the first bisection point, one-half, corresponds to six-tenths of the fractionating range. He now bisects his range of 60 microamperes. If he bisects at 120 microamperes, he has reduced his range to two-tenths. Thus, the second bisection point, one-quarter, corresponds to two-tenths of the fractionating range, etc.

For the first session *Ss* who had found at least five halfway points and whose thresholds remained constant during the scaling procedure were grouped into three classes, depending on the size of their fractionating range. The first class contained those *Ss* with ranges up to 300 microamperes, the second those with ranges from 301–700 microamperes, and the third those with ranges above 700 microamperes.

Within each class, the proportions of fractionating ranges corresponding to each bisection point were totalled and averaged. These were then plotted against the bisection points. For example, suppose there are two *Ss* in any class. For individual *A*, his first bisection point corresponds to five-tenths of his fractionating range. For individual *B*, his first bisection point corresponds to seven-tenths of his fractionating range. Thus, for this class the first bisection point, one-half, corresponds to six-tenths of the fractionating range. This same procedure was carried out for each of the first five halfway points within each class. Thus, for each sex we have three curves, one for each class of fractionating range.

Exactly the same procedure was carried out for the second session's scaling data. In the second session the number of found halfway points was somewhat less than on the first day. Hence, only the first four points were graphed. Because of intraindividual differences in size of fractionating range the points plotted for the two sessions do not always represent the scaling responses of the same *Ss*.

Results and Discussion

Let us first consider the curves for the men. These are in Fig. 2. The curves for the first scaling session show a slight positive acceleration, while the curves for the second session are all negatively accelerated. For both sessions the curve for the mid-size range group shows somewhat less deviation from linearity than do the curves of the two remaining groups. However, in spite of slight differences in slope the three curves for each session closely approximate one another in shape. This indicates that the nature

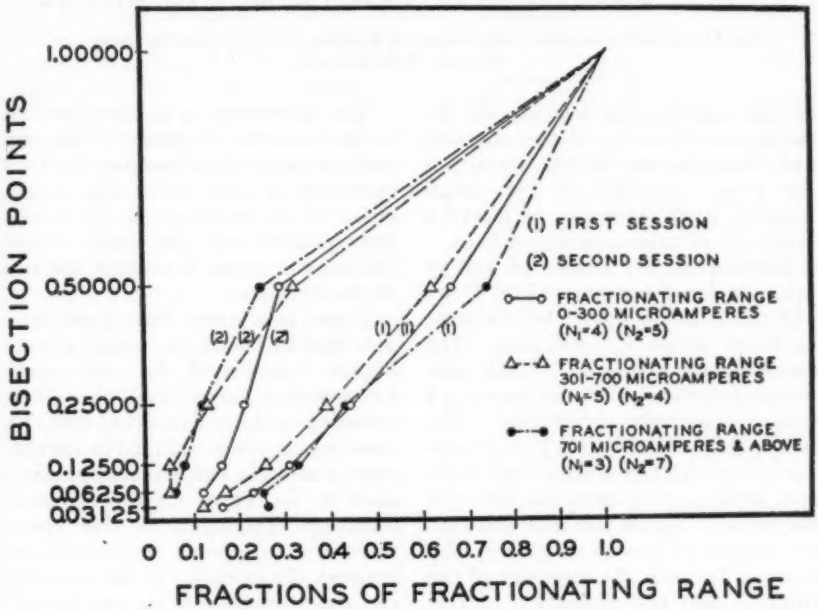

Fig. 2. Successive bisection points plotted as fractions of the fractionating range (men, both sessions)

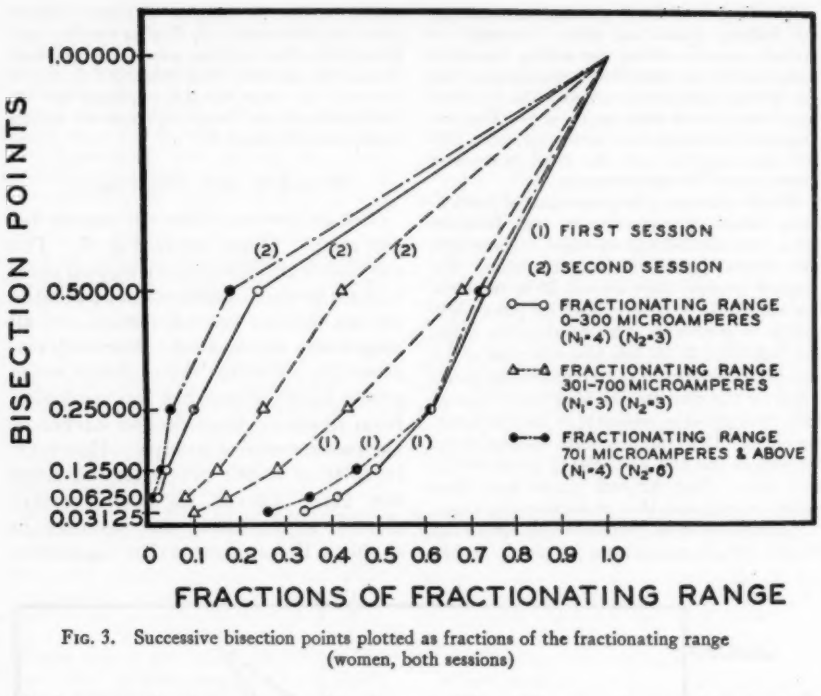

FIG. 3. Successive bisection points plotted as fractions of the fractionating range (women, both sessions)

of the quantitative relationship between the intensity of the stimulus and the estimation of this intensity is the same regardless of the actual intensity of the stimulus, at least for the range of values employed here.

Turning to the curves of scaling behavior for the women (Fig. 3) we find the relationships to be the same as those noted for the men. The curves for the first session are all positively accelerated, those for the second session negatively accelerated. The same pattern of deviation from linearity occurs with the women as with the men, being slightly more marked with the women. Again, this indicates that the nature of the quantitative relationship between the intensity of the stimulus and the estimation of this intensity is the same regardless of the actual value of the stimulus.

The difference in acceleration between the sets of curves for the two sessions probably illustrates the twin operation of time errors due to the order of the limiting stimuli during fractionation and the effect of the direction in which *S* changes the size of the stimulus.

Taken as a whole these facts indicate that had both fractionating techniques been used for the same fractionating range, and the corresponding points averaged, the six resulting curves would tend to approximate a straight line with a slope fairly close to 1. Thus, for the range of intensities employed in this study there is a one-to-one correspondence between the intensity of the stimulus and the estimation of its painfulness.

Superficially, this would seem to corroborate the findings of Hardy,

Wolff, and Goodell (**6, 7**) for low pain intensities. However, careful consideration of the important differences between the studies reveals that a direct comparison may be misleading. First of all, they used the forehead as the site of contact. Secondly, they used radiant heat as a pain stimulus. Thirdly, they employed different units of measurement, reporting data in terms of millicalories per second per square centimeter.

The lack of any scaling-behavior differences between the sexes should not be generalized to cover all forms of the pain response. However, the data are in accord with the results of recent studies which have shown no sex differences in pain-threshold sensitivity (**2, 5, 10**).

Summary

An algesimetric technique using 60-cycle alternating current as the stimulus and the tooth pulp as the site of contact has been described. By this method 24 male and 21 female *S*s successively bisected pain intensities.

For both sexes curves plotted for successive bisections against fractions of fractionating range have the same shape regardless of the actual size of the range being fractionated. Reversing the order of stimulus presentations and the direction in which *S* changes the size of the stimulus, reverses the direction of acceleration of the curves. Compensating for differences in fractionating technique indicates that the relation between the intensity of the stimulus and the estimation of its painfulness approximates a straight line with a slope fairly close to 1.

It is concluded that : (*a*) Within the range of intensities employed in this study there is a one-to-one relationship between the intensity of the stimulus and the estimation of its painfulness; (*b*) there are no sex differences in pain scaling behavior.

(Received September 5, 1952)

References

1. Clark, D., Hough, H., & Wolff, H. G. Experimental studies on headache. *Arch. Neurol. Psychiat., Chicago*, 1936, **35**, 1054–1068.
2. Clausen, J., & King, H. E. Determination of the pain threshold on untrained subjects. *J. Psychol.*, 1950, **30**, 299–306.
3. Goetzl, F. R., Burrill, D. Y., & Ivy, A. C. A critical analysis of algesimetric methods with suggestions for a useful procedure. *Quart. Bull.*, Northwestern Univer. Med. Sch., 1943, **17**, 280–292.
4. Hardy, J. D., & Javert, C. T. Studies on pain: measurement of pain intensity in childbirth. *J. clin. Invest.*, 1949, **28**, 153–162.
5. Hardy, J. D., Wolff, H. G., & Goodell, H. The pain threshold in man. *Amer. J. Psychiat.*, 1943, **99**, 744–751.
6. Hardy, J. D., Wolff, H. G., & Goodell, H. Studies on pain: discrimination of differences in intensity of a pain stimulus as a basis of a scale of pain intensity. *J. clin. Invest.*, 1947, **26**, 1152–1158.
7. Hardy, J. D., Wolff, H. G., & Goodell, H. Studies on pain: an investigation of some quantitative aspects of the dol scale of pain intensity. *J. clin. Invest.*, 1948, **27**, 380–386.
8. Hunter, T. A., & Brown, J. S. A decade-type electronic interval timer. *Amer. J. Psychol.*, 1949, **62**, 570–576.
9. Libman, E. Observations on sensitiveness to pain. *Trans. Ass. Amer. Phys.*, 1926, **41**, 305–308.
10. Schumacher, G. A., Goodell, H., Hardy, J. D., & Wolff, H. G. Uniformity of the pain threshold in man. *Science*, 1940, **92**, 110–112.

Journal of Experimental Psychology
Vol. 45, No. 5, 1953

AN INVESTIGATION OF THE "RANDOMNESS" OF THRESHOLD MEASUREMENTS

MICHAEL WERTHEIMER

Wesleyan University

Ever since the inception of psychophysics, investigators have noticed that successive measurements of a threshold are not identical. Since this variation was generally attributed to experimental error (i.e., errors in the measuring machine and errors made by *E*), a threshold was actually defined in statistical terms, with the assumption that *S* was stable. The result is that measured values are distributed randomly around a central value which, for practical purposes, is then called "the" threshold. Recently, interest has arisen in the temporal course of threshold measurements (**2, 3, 7, 8, 11, 12, 14**) from the point of view that *S* may be unstable. The particular question of interest is whether thresholds do or do not vary randomly in time, and if not, what are the characteristics of their variations. This question, in its more general form, is not limited to sensory thresholds as such, but has important ramifications over the entire realm of biological phenomena, wherever comparable measurements are made.

"Random," as applied to the temporal distributions of sensory thresholds, has at least five possible different meanings: (*a*) that the autocorrelation function of a series of threshold measurements drops to zero at one displacement and never significantly deviates from zero at subsequent displacements; (*b*) that the power spectrum of a series of threshold measurements is not different from the power spectrum of a white noise; (*c*) that a series of threshold measurements in one modality is uncorrelated with a similar series in another modality measured over the same period of time; (*d*) that the power spectrum of a series of threshold measurements in one modality is uncorrelated with the power spectrum of a similar series in another modality measured over the same period of time; and (*e*) that analysis of variance shows that the order of successive threshold measurements on a given *S* does not contribute significantly to the total variance.

These five different definitions of "random" have been examined experimentally in this study. Successive measurements of absolute auditory, visual, and pain thresholds, at intervals of 6 sec., 1 min., 3 min., or 1 day, were performed for a total of 90 series on 35 *S*s in seven different experiments. These measures were then subjected to statistical treatment to see whether they were random according to the different definitions above, and if they were not, to see how their behavior differed from random behavior.

METHOD

The auditory threshold.—For measurement of the auditory absolute threshold, the apparatus consisted of the device used by Hirsh, Rosenblith, and Ward (**6**) for obtaining absolute monaural auditory threshold for clicks, at the anechoic chamber of the Harvard Psycho-Acoustic Laboratory. The auditory stimulus was presented by means of a Colonial pulse generator, type 700A, in the control room. It produced rectangular pulses of 80-msec. duration. Every 3 sec. four clicks were presented, succeeding one another at intervals of .3 sec. These clicks passed through four attenuators so that each pulse was 2 db higher in intensity than its predecessor. All four clicks were then sent through a master attenuator and an impedance matching network, which had a fixed loss of 40

db, to a single permoflux moving-coil earphone (PDR–10) at the right ear of *S*, who sat inside the anechoic chamber. The voltage level of the clicks was adjusted by setting the gain of the pulse generator so that the fourth (most intense) of these clicks produced a peak voltage of 1.3 v. at the input to the impedance matching network when the master attenuator was set at zero. This voltage was determined by producing a deflection on an oscilloscope equal to the zero to peak deflection of a 1000-cycle wave of 5 v. rms. After the initial warm-up period, the intensity of the auditory stimulus was set to the point at which *S* heard four clicks about one half of the time, and less than four the other half. Once this stimulus setting had been made, it remained fixed throughout the rest of the session. In this way it was possible to get a "yes-no" threshold as often as once every 6 sec.

For time intervals longer than 6 sec. the auditory threshold was measured in a slightly different manner. Two successive constant stimuli series of ten responses each were run. Each such series was set at a predetermined attenuation within the range where *S* heard neither exclusively no clicks nor all four clicks.

Finally, in experiments where one day intervened between successive measures, the same apparatus and technique were used except that ten instead of two constant stimuli series were run on each *S*.

The visual threshold.—The apparatus consisted of devices for measuring absolute binocular visual threshold for white light. The Wald adaptometer and the Hecht-Shlaer adaptometer were used, the latter in only one experiment. The Wald (**13**) furnished a red fixation point and a white-light test patch of 2.0° centered 6.0° below the fixation point. The test patch was presented 28 cm. from the eyes for .02 sec. when *S* pressed the shutter button. The brightness of the test patch and the fixation point could be changed together in equal logarithmic steps by means of a rheostat in the control room. Calibration of the test-patch brightness was made with a Macbeth illuminometer; the curve obtained matched the original calibration of the instrument. The shutter, calibrated in H. E. Edgerton's M.I.T. laboratory, proved to have a mean duration of .0212 sec., with a standard error between .0005 and .0006 sec. Repeated settings of the Wald to the same scale value produced changes in brightness of the test patch smaller than the error of measurement of the Macbeth illuminometer.

After at least 5 min. of adaptation to room illumination, Navy goggles for dark adaptation (N-2 type B lens, stock No. 37-L-155-60) were placed on *S*, and an additional 10 min. allowed to elapse. At the end of this period *S* entered the anechoic chamber, the light was turned off, and he was left in total darkness for 20 min. before a record was started. (Most of this period was used in practice and in determination of the auditory threshold.) As with the auditory stimulus, the visual intensity was adjusted to a level at which *S* failed to see the test patch about as often as he succeeded. This stimulus setting, once made, remained fixed through the rest of the session and made it possible to get a "yes-no" threshold as often as once every 6 sec.

For intervals longer than 6 sec. a different technique was used. After the same amount of light and dark adaptation, two pairs of alternately ascending and descending trials were run in a serial method of limits starting with an ascending series. The *S*, after pressing the shutter, signalled either that he did or did not see the test patch. The criterion used was the appearance of two successive "yes" responses in an ascending series, and of two successive "no" responses in a descending series. The time interval between successive presentations of the stimulus was not constant, as *S* released the shutter when ready.

As with the auditory threshold, for experiments in which the interval between measurements was one day, ten instead of two series were run.

In one experiment (Exp. 3) the Hecht-Shlaer adaptometer (**5**) was used instead of the Wald. The *S* was seated at a table inside the anechoic chamber. Before him was a large black cardboard screen out of which projected the eyepiece of Hecht-Shlaer adaptometer Serial No. 6, Model III. The visual stimulus consisted of a test patch of 3.0° visual angle, 7.0° below a red fixation point, presented by a pendulum shutter for .20 sec. at a viewing distance of 25 cm. The light source was calibrated before the experiment so that from a neutral wedge, with a 4 log unit range, the brightness of the test patch could be read off directly in microlamberts. A neutral filter with .1 transmission was also placed between the test patch and the lamp. The room was in total darkness except for a very dull photographic light (with a sharp cutoff, passing only light from the extreme red end of the visible spectrum) which illuminated the controls for *E*. The screen was so arranged that *S* did not report seeing any light whatever even when dark-adapted for an hour and a half. In this technique, the visual threshold was measured by one descending run of serial limits, following an ascending one, each step being .1 log unit. The *E* released the shutter at 4-sec. intervals, and *S* was instructed to respond only if he had been ready when the shutter noise came. A chin rest was used in this experiment.

The pain threshold.—Pain thresholds were

obtained only in experiments with a one-day interval since the pain threshold is strongly affected by the act of measuring (15). The device for measuring the pain threshold was designed after that of Hardy, Wolff, and Goodell (4). The Varitrans Model B-1M was set to read, on the Triplet scale, 70 v. AC, or 275 millicalories/sec./sq. cm. The threshold was measured in terms of the number of seconds (15) that the radiation had to be applied to a circular blackened area, 1.8 cm. in diameter, on the forearm before *S* reported burning pain. This time was measured on a Standard Electric Co. chronoscope, switched on and off at the same time as the radiant energy. For control purposes (16), before measurement of the pain threshold, skin temperature at the blackened circle was taken by means of a thermistor (1V611) and impedance bridge, Type 650, Serial No. 2, General Radio Co. Resistance was translated into degrees centigrade by means of a calibration curve.

Other equipment, and subjects.—Devices were available in all experiments to permit *S* and *E* to communicate between experimental room and control chamber automatically. The *S*s were all from a college population, including both men and women, and undergraduates, graduate students, and faculty members.

Experimental Design

The experimental design called for seven different experiments.

Experiment 1a.—*Auditory* and *visual* thresholds on three *S*s were measured alternately, so that a pair of measurements was obtained at 6-sec. intervals for about one hour. This procedure was repeated on three different occasions, giving a total of 9 records for the three *S*s. Ten other *S*s were used for about one hour, each of them one time only. Recording of responses was automatically performed by an Esterline-Angus kymograph for the last 17 records, by hand for the first 2. A total of 19 records was obtained.

Experiment 1b.—The data from Exp. 1*a* were transformed into threshold measurements of per cent of presented stimuli responded to as sensed. By this conversion, pairs of measurements were obtained at *1-min.* intervals. A total of 19 records was available.

Experiment 1c.—The data from Exp. 1*a* were converted as in Exp. 1*b*, except that pairs of measurements were made available at *3-min.* intervals. Fourteen records were obtained which included 15 or more pairs.

Experiment 2.—*Auditory* and *visual* thresholds were measured at *3-min.* intervals by the method of constant stimuli for about one hour on three *S*s, each of these *S*s being used three times, giving a total of 9 records. Ten other *S*s were used for about one hour each. A total of 19 records was thus obtained, 16 of them having 15 or more pairs.

Experiment 3.—Using the Hecht-Shlaer adaptometer, ten *S*s were each used for 45 min., taking *auditory* and *visual* thresholds by the method of constant stimuli at *3-min.* intervals. A total of ten records was obtained in this experiment.

Experiment 4a.—On two *S*s, *auditory* thresholds alone were taken at *1-day* intervals for 25 consecutive days, producing a total of two records.

Experiment 4b.—*Auditory* and *pain* thresholds were taken at *1-day* intervals for 11 consecutive days on three *S*s, producing a total of three records.

Experiment 4c.—*Auditory* and *pain* thresholds were taken at *1-day* intervals for 45 consecutive days on three *S*s, producing a total of three records.

Experiment 4d.—*Auditory* and *visual* thresholds were taken at *1-day* intervals for 23 consecutive days on three *S*s, producing a total of three records.

Methods of Analysis

1. The autocorrelation functions of the records were calculated. Each series of measurements was correlated with itself displaced by one, two, etc., measures, up to ten. In Exp. 1*a* a tetrachoric correlation coefficient was estimated (1); in the remainder a Pearson product-moment correlation coefficient was calculated at each lag.

2. For every threshold series, the autocorrelation function was transformed into an estimated power spectrum, according to Tukey's technique (9, 10). This power-spectrum analysis treats the successive threshold measures as though they were successive samples of the amplitude of a "noise." From the completed analysis one can make crude inferences from an observed sample to the spectrum of the noise that produced it. Using the serial product at successive lags or displacements, this technique permits one to extract information about the "color" of the noise. The lag coefficients (Q_p) are multiplied successively by a matrix which yields "raw" estimates of the power spectrum (L_h) and by a matrix which serves to give "smoothed" estimates of the power spectrum (U_h).

Beginning with the observed data x_n ($n = 1, 2, \ldots, N$) the lag coefficients are calculated (as in the autocorrelations above):

$$Q_p = \frac{1}{N-p} \sum_{n=1}^{N-p} (x_n)(x_{n+p}), \quad (p = 0, 1, \ldots, m).$$

Next the "rough" estimates of the power spectrum:

$$L_h = \frac{1}{m}\left(Q_0 + 2\sum_{p=1}^{m-1} Q_p \cos\frac{\pi p h}{m} + Q_m \cos \pi h\right),$$
$$(h = 0, 1, \ldots, m).$$

Finally, the "smoothed" estimates of the power spectrum are calculated:

$$U_h = .23L_{h-1} + .54L_h + .23L_{h+1},$$

where $L_{-1} = L_1$ and $L_{m+1} = L_{m-1}$.

The results of such an analysis permit one to draw a curve showing relative amounts of power in the noise falling at frequencies corresponding to the lag used. The amount of information that can be derived from such a plot depends upon the size of the sample, the width of the "filter," and the "distance" between successive places on the frequency continuum at which the filter is applied. With the smoothing function above, the width of the filter becomes very large relative to the distance moved along the continuum, resulting only in quite gross estimates of relative amounts of power present in different bands. It also goes without saying, of course, that conclusions about the color of the noise can only be made within the limits of frequencies tested (i.e., the abscissae used) and the very coarse filter used (i.e., the number of places on the abscissa at which the filter is applied, and the extent of smoothing).

3. Each series of threshold measurements taken over a given period of time was correlated with the other series of thresholds taken on the same *S* over the same period of time, using a Pearson product-moment correlation coefficient, except in the case of the 6-sec. thresholds (Exp. 1a), where a tetrachoric correlation coefficient (1) was employed.

4. Using Spearman's ranks correlation, the power spectrum of every threshold series was correlated with the power spectrum of the other series of threshold measures, taken in a different modality over the same period of time on a given *S*.

5. On the measurements obtained with intervals of one day, analysis of variance was performed with three variables: days, *S*s, and order of threshold measurement.

6. Throughout, nonparametric techniques were employed in preference to standard ones, to assess statistical significance of findings.

Results

Autocorrelation.—For all time intervals, thresholds showed significant deviation of the autocorrelation functions from chance. Entered in Table 1 are the number of records in a given experiment which showed a positive correlation at the successive displacements, and the probability of obtaining this many (or this few) by chance. Dashes are entered in the table when the nonparametric probability of obtaining *N* positive correlations by change is greater than .10. Since the number of records on daily visual and

TABLE 1

Nonparametric Autocorrelation Functions

Exp.	Time Interval	Number of Records		Displacements: Auditory Series									Displacements: Visual Series								
				1	2	3	4	5	6	7	8	9	1	2	3	4	5	6	7	8	9
1a	6 sec.	17	*N*	17	10	9	8	7	6	5	6	6	17	15	13	13	12	10	7	5	7
			p	.01	—	—	—	—	—	.07	—	—	.01	.01	.02	.02	.07	—	—	.07	—
1b	1 min.	19	*N*	14	14	13	10	10	10	10	7	12	14	11	10	8	5	6	7	10	12
			p	.03	.03	.08	—	—	—	—	—	—	.03	—	—	—	.03	.08	—	—	—
1c	3 min.	14	*N*	8	8	7	6	8	6	5	9	10	9	6	8	9	6	8	8	7	5
			p	—	—	—	—	—	—	—	—	.09	—	—	—	—	—	—	—	—	—
2	3 min.	16	*N*	12	10	8	9	9	4	5	6	9	14	10	7	10	10	6	7	6	7
			p	.02	—	—	—	—	.02	.06	—	—	.01	—	—	—	—	—	—	—	—
3	3 min.	10	*N*	9	7	7	8	7	4	7	8	3	3	4	5	6	4	2	3	7	1
			p	.01	—	—	.05	—	—	—	.05	—	—	—	—	—	—	.05	—	—	.01
1c, 2, & 3 combined	3 min.	40	*N*	29	25	22	23	24	14	17	23	22	26	20	20	25	20	16	18	20	13
			p	.01	.08	—	—	—	.04	—	—	—	.04	—	—	.08	—	—	—	—	.02
4	1 day	11	*N*	11	6	3	4	5	7	6	2	6									
			p	.01	—	—	—	—	—	—	.03	—									

pain measures was too small, their values are not entered. The pain N's are 6, 5, 2, 2, 4, 2, 3, 4, 1, 3.

These data suggest that successive measurements of a threshold are not independent, regardless of which time interval is used. Parametric tests, it should be mentioned, gave even greater significance levels than did the nonparametric ones in the table (**14**).

Power spectra.—The average estimated power spectrum transforms obtained in all experiments are shown in Fig. 1, 2, and 3. Figures 1 and 2 show the mean estimated power spectra from periods of 10 sec. through 27 min., for the auditory and visual thresholds. In both figures, the five sets of data have been plotted so as to give a suggestive picture; the position of the five sets of data relative to each other is arbitrary, but the position of the eight points from any one experiment relative to one another is in absolute obtained terms. Both figures show more activity in the slower periods than in the faster, and both deviate significantly from the straight line which would be obtained if the spectra had not been different from that of a white noise.

Figure 3 shows the mean estimated power spectra from periods of two through nine days, for all three modalities studied, as well as gives some suggestion of the reliability of such spectra. It should be mentioned that the apparent nine-day peak in the visual threshold spectrum is believed to be artifactual (**14**).

Table 2 shows the nonparametric power spectra. These were prepared in the following way: of the eight points calculated in the spectrum, the four highest are rated 1 and the four lowest 0. Entered in the table are the number of 1's at each period, and the probability of obtaining this many (or

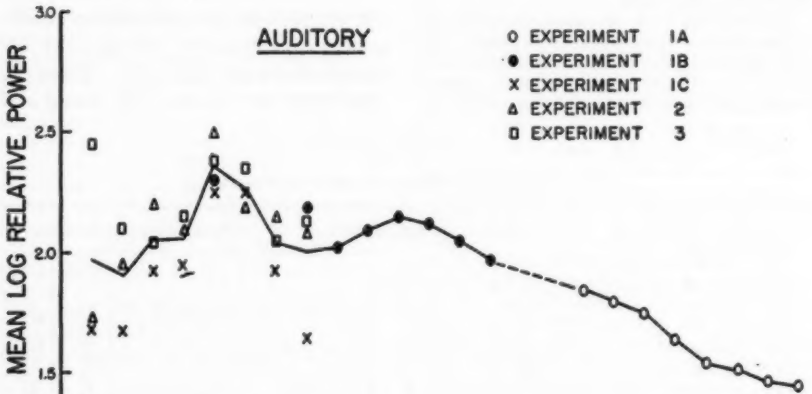

FIG. 1. Mean estimated power spectrum of auditory threshold measurements, for periods of 10 sec. to 27 min.

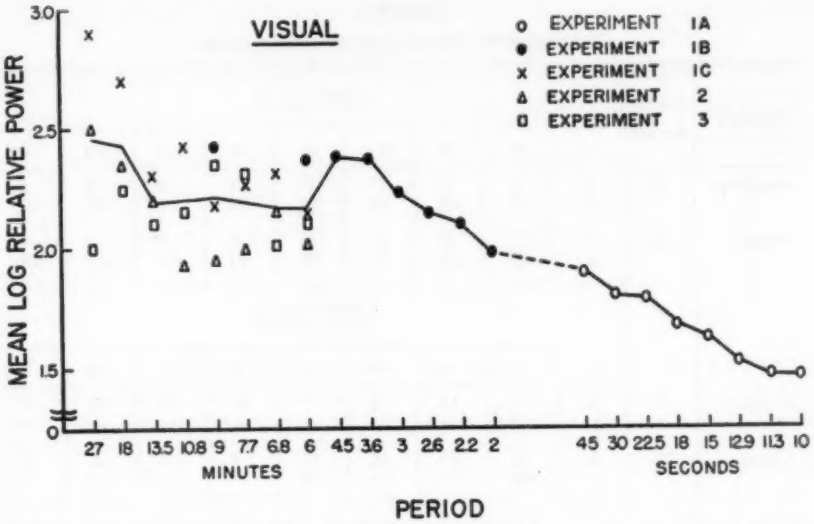

FIG. 2. Mean estimated power spectrum of visual threshold measurements, for periods of 10 sec. to 27 min.

this few) by chance. Dashes are entered in the table when the non-parametric probability of obtaining N 1's by chance is greater than .10; once more it should be mentioned that parametric tests (14) gave far greater significance levels than did the non-parametric ones entered in the table. It will be noticed that the data from Exp. 4 are missing from the table; since the number of records on daily measurements was too small, their values were left out. The significance of cyclical activity over periods of days can be estimated from Fig. 3, where some standard errors are given.

As can be seen from Fig. 1, 2, and 3, and Table 2, the power spectra are significantly different from those obtained in a series which is not different from white noise. In fact, there are clear indications of periodicities in the different modalities. A highly repeatable 9-min. cycle was evident in the auditory threshold, and there were suggestions of both an 18-min. and a 6-day auditory cycle; a 2-day pain cycle; and visual cycles of 4 min., 8 min., and 2 days. Auditory

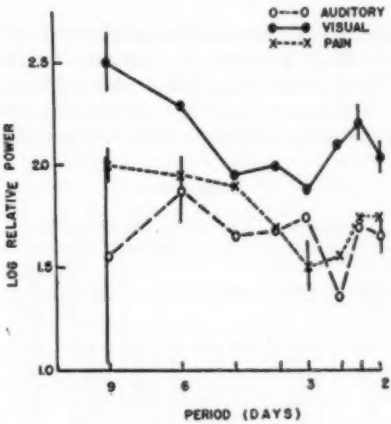

FIG. 3. Mean estimated power spectra of auditory, visual, and pain thresholds, for periods of two to nine days. The vertical lines show the standard error of some representative points.

TABLE 2

NONPARAMETRIC POWER SPECTRUM TRANSFORMS

Modality	Number of Records		Experiment 1a Period (seconds)							
			45	30	22.5	18	15	12.9	11.3	10
Auditory	17	*N*	17	17	15	17	2	0	0	0
		p	.01	.01	.01	.01	.01	.01	.01	.01
Visual	17	*N*	17	17	17	15	0	0	1	1
		p	.01	.01	.01	.01	.01	.01	.01	.01
			Experiment 1b Period (minutes)							
			9	6	4.5	3.6	3	2.6	2.2	2
Auditory	19	*N*	11	13	8	7	11	9	10	7
		p	.08	—	—	—	—	—	—	—
Visual	19	*N*	14	10	12	17	8	5	6	4
		p	.03	—	—	.01	—	—	—	—
			Experiments 1c, 2, & 3 combined Period (minutes)							
			27	18	13.5	10.8	9	7.7	6.8	6
Auditory	40	*N*	18	14	18	23	26	26	20	17
		p	—	.02	—	—	.02	.02	—	—
Visual	40	*N*	24	24	23	19	16	14	20	20
		p	.08	.08	—	—	.08	.02	—	—

and visual thresholds also showed less random activity in periods below about 2 min. than below about 20 sec.

Correlations.—All experiments showed small correlations between the thresholds of different modalities, although not always significant alone. These correlations are shown in Table 3. Auditory and visual thresholds are positively correlated, and auditory and pain thresholds negatively. Considering solely the number of positive and negative correlations obtained, the first of the two preceding statements has a probability of being incorrect of less than .001; the second has a similar probability of .02. Unfortunately the design did not make it possible to test for the possible correlation between visual and pain thresholds.

TABLE 3

INTERMODAL THRESHOLD CORRELATIONS

Exp. Number	Time Interval	Modalities	*N*	Mean *r*	Number of Positive *r*'s
1a	6 sec.	Aud. and vis.	19	+.25	15
1b	1 min.	Aud. and vis.	19	+.17	14
1c	3 min.	Aud. and vis.	14	+.23	9
2	3 min.	Aud. and vis.	19	+.27	13
3	3 min.	Aud. and vis.	10	−.02	6
4b	1 day	Aud. and pain	3	−.10	0
4c	1 day	Aud. and pain	3	−.14	0
4d	1 day	Aud. and vis.	3	+.21	3
Total		Aud. and vis.	84	+.21	60
Total		Aud. and pain	6	−.12	0

Correlation of power spectra.—The threshold spectra of different modalities taken over the same time interval

TABLE 4

RANK CORRELATIONS OF POWER SPECTRA

Exp. Number	Time Interval	Modalities	*N*	*rho* of Mean Spectra	Mean *rho*	Number of *rhos* More Positive than *rho* of Mean Spectra
1*a*	6 sec.	Aud. and vis.	17	+.80	+.93	10
1*b*	1 min.	Aud. and vis.	19	+.50	+.13	8
1*c*	3 min.	Aud. and vis.	12	−.35	+.25	10
2	3 min.	Aud. and vis.	16	−.38	+.05	13
3	3 min.	Aud. and vis.	10	+.09	+.40	8
Total		Aud. and vis.	74	+.15	+.35	49

on the same *S* tended, by Spearman ranks correlation, to be more similar than were the average spectra of the same modalities' thresholds on different *S*s (Table 4). In the table, *rho* of the Mean Spectra refers to the correlation obtained in each experiment between the final mean spectra of each modality investigated in that experiment. The Mean *rho* is the average correlation obtained when the spectra of *each record* in an experiment were compared. No entries are made in the table for the spectra of Exp. 4, since the number of such correlations was too few, and each correlation is too unreliable, since the number of pairs for each correlation was only eight.

The data suggest that the threshold variations of two modalities show some kind of similar pattern of change over a given time period, whether they correlate positively or not.

Analysis of variance.—Analysis of variance of the auditory and visual thresholds obtained in Exp. 4*d* showed that the day-by-subject interaction contributed more (at the .1% level of significance) to the total variance than did the error term, showing that the thresholds were not varying randomly from day to day in a given *S*, but that there were some significant changes, different for the different *S*s (since days did not contribute significantly to the total variance, indicating also that the drift was not due to the apparatus, weather, etc.), in the thresholds over the time during which they were measured. The *F*'s were 25.94 and 49.38 for the subject-by-day interactions, tested against the triple-interaction error term, for the auditory and visual thresholds respectively.

DISCUSSION

Indications in the experiments were that the longer the time samples one takes, the larger will be the *SD* of a threshold, suggesting that there are very slow major changes in the thresholds over time (**14**). These changes of threshold from day to day are statistically reliable, showing that the thresholds undergo drifts not attributable to "purely random" fluctuation or to mechanical instability of the equipment alone. The correlation between thresholds of two different modalities suggests some kind of interdependence of the senses measured. The nonindependence of successive measures indicates that there is some kind of a damping factor in the sensory mechanism, so that successive measures are not random. The fact that the power spectrum transforms of these autocorrelation functions

show definite cycles, suggests that some kind of physiological rhythms are being exhibited in the sensory phenomena investigated. What the characteristics of these physiological factors are, is impossible to discern from these experiments; the only hint is that the physiological factors must be rhythmically affecting different sensory systems at the same time, since the power spectra of the thresholds of different modalities tend to be related, in that the *patterns* of temporal drift of the thresholds of two modalities of an *S* are similar over the same period of time.

Considering the interrelations of the threshold measures, it would seem that a view of sensory functioning which assumes a functional isolation among the senses would be too parsimonious. A view which holds that the senses are integral parts of the same total organism and cannot be assumed to be functionally isolated distinct units, that each sense is a functioning part of the total organism, would seem to have at least heuristic value. Assumptions concerning the independence or randomness of such organismic functions do not seem warranted.

Summary and Conclusions

Successive measurements of auditory, visual, and pain thresholds were obtained at 6-sec., 1-min., 3-min., and 1-day intervals on a series of *S*s. These psychophysical thresholds were subjected to five different kinds of analysis: autocorrelation, estimated power-spectrum transformation of the autocorrelation function, correlation of the thresholds of two modalities measured over the same period of time on the same *S*, correlation of the spectra of the thresholds of two modalities taken over the same period of time on the same *S*, and analysis of variance.

The data obtained tended to confirm the hypothesis that the variations of thresholds in time, under normal conditions, are not "random." This evidence is: (*a*) Reliable changes in thresholds can be shown to occur in *S* from day to day, (*b*) changes in the threshold of different modalities are slightly correlated, (*c*) successive threshold measurements are not independent, (*d*) some thresholds can be shown to undergo cyclical drifts, and (*e*) the patterns of variation of thresholds of different modalities over the same time interval in the same *S* are not unrelated. The relevance of these findings for an approach to the total organism is discussed.

(Received September 2, 1952)

References

1. Cheshire, L., Saffir, M., & Thurstone, L. L. *Computing diagrams for the tetrachoric correlation coefficient.* Chicago: Univer. of Chicago Bookstore, 1933.
2. Collier, G. H. An investigation of the Poisson frequency of seeing function and of some of the possible variables governing the independence of successive responses at the visual threshold. Unpublished doctor's thesis, Indiana Univer., 1951.
3. Day, W. F. Stimulus interval as a determiner of serial patterns of threshold responses. Unpublished master's thesis, Univer. of Virginia, 1951.
4. Hardy, J. D., Wolff, H. G., & Goodell, H. Studies on pain. A new method for measuring pain threshold; observations on spatial summation of pain. *J. clin. Invest.*, 1940, **19**, 649–658.
5. Hecht, S., & Shlaer, S. An adaptometer for measuring human dark adaptation. *J. opt. Soc. Amer.*, 1938, **28**, 269–275.
6. Hirsh, I. J., Rosenblith, W., & Ward, W. D. The masking of clicks by pure tones and bands of noise. *J. acoust. Soc. Amer.*, 1950, **22**, 631–637.
7. Lee, R. H., Finch, E. M., & Pounds, G. A. Periodic fluctuations in the dark adapted threshold. *Amer. J. Physiol.*, 1945, **143**, 6–10.

8. Myers, C. K., & Harris, J. D. Variability of the auditory threshold with time. Medical Research Lab. (New London) Report No. 165, 1950, 230–257.
9. Tukey, J. W. *Measuring noise color.* Bell Telephone Laboratories, Inc., Case 20878 (undated).
10. Tukey, J. W. *The sampling theory of power spectrum estimates.* Bell Telephone Laboratories, Inc., December, 1949.
11. Verplanck, W. S. An experimental test of Hecht's quantum theory of the probability of seeing function. *Amer. Psychologist,* 1948, **3,** 358. (Abstract.)
12. Verplanck, W. S., Collier, G. H., & Cotton, J. W. Nonindependence of successive responses in measurements of the visual threshold. *J. exp. Psychol.,* 1952, **44,** 273–282.
13. Wald, G. A portable visual adaptometer. *J. opt. Soc. Amer.,* 1941, **31,** 235–238.
14. Wertheimer, M. A study of normal threshold variations in time. Unpublished doctor's thesis, Harvard Univer., 1951.
15. Wertheimer, M. A temporal method for measurement of the pain threshold. *Amer. J. Psychol.,* 1952, **65,** 297–298.
16. Wertheimer, M., & Ward, W. D. The influence of skin temperature upon the pain threshold as evoked by thermal radiation —a confirmation. *Science,* 1952, **115,** 499–500.

Journal of Experimental Psychology
Vol. 45, No. 5, 1953

FOVEAL SIMULTANEOUS BRIGHTNESS CONTRAST AS A FUNCTION OF INDUCING- AND TEST-FIELD LUMINANCES[1]

A. L. DIAMOND

Columbia University

A small lighted square will appear to change in brightness if another lighted square, of different brightness, is presented beside it. This is the phenomenon of simultaneous brightness contrast. The square on which we discriminate the "apparent" brightness[2] change may be termed the test field. The adjacent stimulus may be called the inducing field.

Previous *E*s have found, generally, that as the luminance of the inducing field increases, the apparent brightness of the test field decreases (**4, 8, 9, 10,** p. 567–571).

The experiment of Hess and Pretori (**8**) employed a monocular matching technique involving four visual fields. That is, a 10° inducing-field square (10° on a side in visual angle), centered about a 1° test-field square, was adjacent to another 10° surround-field square centered about a 1° match-field square. The *S*'s task was to indicate an apparent-brightness match between the test and match fields for different luminance conditions of the four fields. It was found, in general, that the apparent brightness of the test field decreased in direct proportion to the luminance of the inducing field. This result held true for several different ratios of match-field luminance to the luminance of its surround.

Hess and Pretori's experiment illustrates the contrast effect, i.e., the inhibition of the test-field apparent brightness as a direct function of the luminance of the inducing field. The precise details of their data are, however, difficult to evaluate, since they used four fields. For this reason, the contrast effect was not only determined by the effect of the inducing field upon the test field but also by the interacting effects of (*a*) the match field upon the test field, (*b*) the match field upon the inducing field, (*c*) the match-field surround upon the test field, and (*d*) the match-field surround upon the inducing field.

In a contrast experiment by Fry and Alpern (**4**) this objection is met by a technique involving a binocular apparent-brightness match; that is, only the test field (a 2.5° × .5° illuminated rectangle) and the inducing fields (two "glare source" fields equal in size and shape to the test field and placed symmetrically on either side of it) are presented to the right eye of *S*. The match field (also a 2.5° × .5° rectangle) is presented to the left eye of *S* thus making insignificant the effect of any interaction between the match field and the test and inducing fields. Fry and Alpern have shown binocular effects to be small compared to the monocular contrast effects.

One of the parameters investigated by Fry and Alpern was the luminance of the inducing fields over a range of

[1] This work was supported by a contract between the Office of Naval Research, Department of the Navy, and Columbia University. The writer is indebted to Professor Clarence H. Graham and Professor Conrad G. Mueller for their guidance and encouragement. The writer is now at Northwestern University.

[2] The terms *apparent brightness* and *brightness* are used synonomously in the text to indicate the subjective aspect of *S*'s response.

from about 6 ft.-lamberts (the luminance of the test field) to about 140,000 ft.-lamberts. The results, generally similar to those of Hess and Pretori, show that the apparent brightness of the test field is depressed as the inducing-field luminance is increased.

In the Fry and Alpern study, the inducing-field luminance was never much less than that of the test field. Also, the test-field luminance was not varied systematically as a parameter of the entire contrast function. In the present experiment, therefore, these two factors are of especial interest. The luminance of the inducing field is varied from considerably below that of the test field to a value much higher than that of the test field. In addition, the test-field luminance is also varied over a relatively wide range.

The specific design of the present experiment is as follows: The inducing and the test fields are adjacent squares equal in size and presented to *S*'s right eye (see Fig. 1A). A match field (*m*) equal in size to the test field (*t*) is presented to *S*'s left eye. The contrast effect of the inducing-field (*i*) luminance upon the test-field luminance is thus indicated in the equality brightness match between the match and the test fields. In this manner, the apparent brightness of the test field is measured as a function of the inducing-field luminance. This function is studied in two related experiments: (I) for seven test-field luminance levels; and (II) for three apparent binocular separations between the match field and the test and inducing fields.

Apparatus

The apparatus used in both experiments is essentially the same. Minor modifications are considered in the discussions of the respective experiments.

The view presented to *S* is shown in Fig. 1A. The apparatus provides as a measure of the apparent brightness of the test field (*t*) the luminance of the match field (*m*), which can be set equal to the apparent brightness of the test field by *S*. Pattern *R* is presented to *S*'s right eye; it includes the inducing field (*i*), a square subtending 33′ of visual angle on a side, and the test field (*t*), equal in size to, and immediately below, the inducing field. The *S*'s right eye fixates upon a 3′ point of light (P_2) located 21′ to the left of

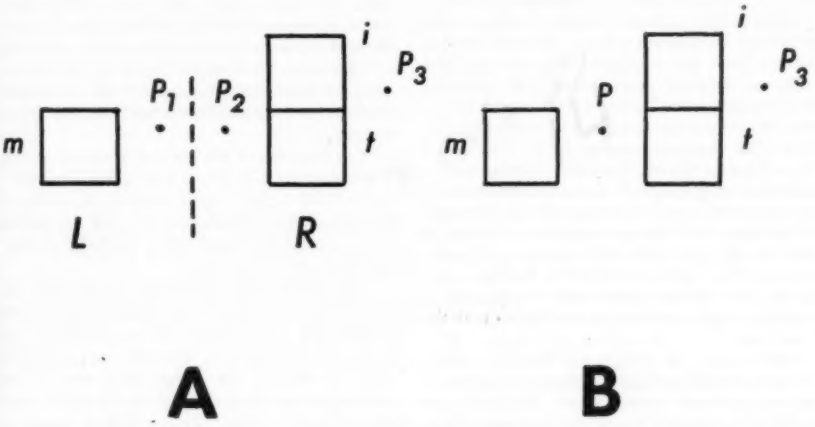

Fig. 1. (A) Patterns *L* and *R* as seen monocularly by *S*; *L* is presented to the left eye only, *R* to the right eye only. (B) The binocular view when fixation points P_1 and P_2 are fused.

the test field and 21′ above an extension of the bottom line of the test square. Both fields are adjustable in luminance by *E*. Pattern *L* is presented to *S*'s left eye; it includes the match field (*m*), equal in size to the test field and adjustable in luminance by *S*. The *S*'s left eye fixates a 3′ point of light (P_1) located 21′ to the right of match field (*m*) and 21′ above an extension of the bottom line of the match square. The *S* is instructed to fuse left- and right-eye fixation points so that in the binocular view shown in Fig. 1B, only one fixation point (*P*) is seen between the match and the test fields, and the two fields are thus held by fusion of the fixation points in constant position relative to one another.[3] Stimulation may be considered essentially foveal since the visual angle between the fused fixation point and the farthest corner of any field is 70′.

The visual pattern in Fig. 1A is presented to *S* along two separate optical paths, one to each eye. The light source of each path is a 150-w. tungsten filament projection lamp, in front of which is a section of heat-absorbing glass. The light, diffused by flashed opal glass, travels through its particular pattern, *L* or *R*. The luminance can be continuously varied by fixed filters and a stationary and movable polaroid. Next, made parallel by a 4-diopter lens, the light finally travels through a 3-mm. artificial pupil into the eye of *S* who is seated in a light-tight cubicle. The entire left optical path is adjustable horizontally for interpupillary distance by means of a screw arrangement attached to the optical bench.

Consider now the right optical path, specifically with reference to control of the inducing- and test-field luminances. Ten pattern *R*'s (see Fig. 1A) are cut in a thin metal disc which can be rotated and set so that each pattern falls directly before the diffuse light source. Neutral Wratten gelatin filters are mounted behind the inducing fields of eight of these patterns; the inducing field of one pattern is of zero density and that at the other extreme is completely blacked out. The filter densities increase from .5 to 4.2 in approximate steps of .5 log unit. The luminance of the inducing field may thus be decreased from its maximum to zero luminance as the luminance of the test field remains constant. Pattern *R* can be optically inverted, moreover, by the rotation of a Dove prism set in the right eyepiece. This device provides for the squares in pattern *R* to be reversed in position so that the square behind which the gelatin filter is mounted now appears as the test field. The test-field luminance may thus be varied while the inducing-field luminance remains constant. For the field behind which the gelatin filter does not appear (depending upon the position of the Dove prism) the luminance may be varied by the fixed filters and the variable polaroid. (These filters reduce the luminance of both fields simultaneously so that a desired luminance of the field behind which the gelatin filter does appear is calculated from the combination of the densities of the gelatin filter, the fixed filters, and the polaroid.) Within the luminance limits available, therefore, any desired combination of inducing- and test-field luminances can be obtained through combined manipulation of the thin metal disc, the Dove prism, the fixed filters, and the variable polaroid.

It should be noted that, as pattern *R* is inverted by the Dove prism, the square which was formerly the inducing field becomes the test field. This fact necessitates fixation point P_3 which serves in place of P_2 when the squares are reversed in position.

The luminance of the match field (*m*) in the left optical path, is controlled in the following manner: Between the diffuse light source and pattern *L* is a small polaroid disc which just covers the match field but not its fixation point. This disc serves as the fixed member of a pair of polaroids; the variable member is coupled to a metal disc which *S* can rotate about the left eyepiece. As *S* varies the brightness of the match field, therefore, the fixation-point luminance remains constant. The luminances of both the match field and its fixation point are simultaneously changed by fixed filters. The luminance of the match field is indicated by the combined densities of the fixed filters and polaroids.

Apparatus calibrations involved determinations of filter densities and of the maximum luminance of the match, the test, and the inducing fields.

The densities of the neutral Wratten gelatin filters and the fixed filters were determined photometrically as was transmission of the polaroid in various angular positions. During the experiment proper, no value of a polaroid reading greater than 75° was used.

The absolute luminance of the match field was determined by a binocular matching procedure (identical to that followed in the experiments to be described) that involved another square field, equal in size to the match field, and whose luminance had been determined by a Macbeth illuminometer. This determination gave a figure of maximum available match-field luminance of 1178 ml. The maximum available

[3] The task of fusing the fixation points requires a small amount of training in control of convergence and becomes more difficult as the luminance of the fixation points decreases. Since in the present experiment both *S*s were orthophoric, very little training was necessary.

luminances of the test field and inducing field were also 1178 ml.[4]

It should be noted that in the experimental procedure, the luminances of the test and inducing fields never exceeded 513 ml. Since *S* matched the apparent brightnesses of the test and inducing fields according to the psychophysical method of adjustment, this enabled him to vary the match-field luminance above, as well as below, that of the test field, even when the test-field luminance was at its maximum.

Experiment I

Method.—In this experiment, the apparent brightness of the test field is studied as a function of the inducing-field luminance over seven test-field luminance levels; that is, *S* matches the match field (*m*) to the test field (*t*) (in Fig. 1B) at each of eight inducing-field (*i*) luminances. This procedure is repeated for seven luminance levels of the test field ranging from −.39 to 2.71 log ml. Each experimental session is thus devoted to obtaining the function for a constant luminance of the test field.

The procedure involved in one experimental session required approximately 40 min. Seated in a light-tight cubicle, *S* initially dark adapted for 10 min. if the test-field luminance used in a particular session was less than 10 ml., the approximate average luminance to which *S* was light adapted before entering the cubicle. There was no initial dark adaptation for test-field luminances greater than 10 ml. The *S* then viewed the pattern shown in Fig. 1B and light adapted to (*a*) the test field, (*b*) the inducing field at minimum luminance, and (*c*) the match field (at the same luminance as the test field) in the other eye. Light adaptation was for 1.5 min. during which time four (or sometimes five) "warm-up" brightness-equality matches were made of the match and the test fields. The procedure generally followed thus avoided prolonged light adaptation to high test-field luminances and was found to be effective in controlling fatigue and increasing the consistency and accuracy of the data. The data under this procedure are comparable to those obtained under other conditions of adaptation described later.

In making brightness-equality matches *S* followed the psychophysical method of adjustment as described by Guilford (6), i.e., he set the luminance of the match field so that it appeared equal to that of the test field. The *E* then changed the luminance of the match field in a random manner after each match. The *S* had to adjust the match-field luminance again until it seemed equal in apparent brightness to that of the test field. In this manner *S* made a group of ten equality matches at each of eight inducing-field luminances (−1.51 to 2.71 log ml.), the initial luminance being near or below threshold. The inducing-field luminance was increased after each set of ten matches. Dark adaptation for .5 min. followed each group of ten matches at each inducing-field luminance; 1.5 min. of light adaptation and at least four "warm-up" matches preceded the next group of ten matches at the next inducing-field luminance. The average of each group of ten match-field luminances was taken as the measure of the apparent brightness of the test field at each value of the inducing-field luminance. This procedure was followed for the seven test-field luminance levels and for two *S*s.

A control was further performed to check the findings of the above procedure. Since dark adaptation was omitted in the previously described determination at the high test-field luminance levels, the retinal area upon which the inducing field fell was relatively light adapted during the first 6 min. of the procedure, that is, while *S* was making matches at the low luminance values of the inducing field. The results are tested, therefore, by the following procedure in which influences due to adaptation effects are controlled.

Each experimental session produced data which are distributed among the seven curves of Fig. 2. At the beginning of each session *S* dark adapted 10 min. and then light adapted to the lowest test-field and inducing-field luminances for that session. The luminances of both the test and inducing fields were then equally increased in steps of approximately .5 log ml. As in the original procedure above, .5 min. of dark adaptation, 1.5 min. of light adaptation, and four practice matches preceded ten matches at each step. Four such sessions were conducted for one *S*, A. D., the sessions beginning at test-field luminances of, respectively, 1.15, −.39, −.39, and −.39 log ml. The second and fourth of these sessions were repeated and the data of each repetition were averaged with the original data (see daggered entries in Table 1). The procedure as outlined, therefore, provided that at the beginning of each session the inducing-field retinal area was completely dark adapted, and that at each successive group of matches the light adaptation necessary for both the test and the inducing fields was essentially the same.

One further extension of the data involved a brightness match between the match and the test fields without the inducing field present. The foregoing procedure was followed except

[4] This is actually a unit of luminous emittance according to Chapanis (3, p. 47–49), although for convenience the term "luminance" is used in the present experiment.

that the inducing field was covered. This was done in one session for both *Ss*, A. D. and H. L.

Results.—The results are presented in Table 1 and in Fig. 2. Table 1 presents the individual results for the two *Ss*. Figure 2 shows the luminance of the match field (B_m) for the judgment of equality, plotted as a function of the inducing-field luminance (B_i). B_m may be termed the "apparent" brightness of the test field. The function is graphed for different levels of the the test-field luminance. It should be noted that the curves drawn through the points are theoretical, as will be discussed later.

Figure 2 shows that the data of the control procedure (plotted as open circles and triangles) are quite consistent with the data of the original procedure (plotted as closed circles and triangles). The shape of the contrast function for the fovea, therefore, seems to be accurately represented by the curves of Fig. 2.

These curves, generally, show that the test-field apparent brightness, as measured by the match-field luminance, decreases relatively little as the inducing-field luminance is increased from zero to about equal to that of the test field. For inducing-field luminances greater than that of the test field, the test-field apparent brightness decreases approximately in direct proportion to the luminance of the inducing field. Thus, there is a break or a bend in the function which occurs when the inducing- and test-field luminances are equal. This bend, moreover, moves towards the right along the inducing-field luminance axis as the test-field luminance, for a particular curve, is increased. That the function in Fig. 2 is fitted with a continuous line rather than two straight lines on either side of the break or bend is due to a number of theoretical considerations which will be discussed later.

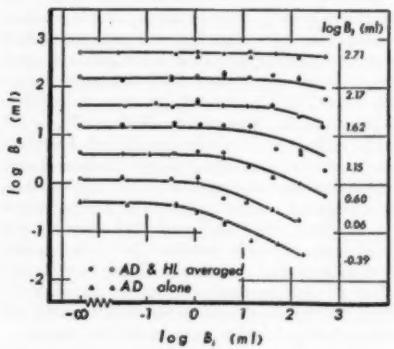

FIG. 2. Apparent brightness of the test field (B_m) as a function of inducing-field luminance (B_i), for different test-field luminances (B_t)

EXPERIMENT II

This experiment is concerned with a possible interaction effect between the inducing field in one eye and the match field in the other. If such an effect does exist, the match field as well as the test field changes in apparent brightness as the inducing-field luminance varies; our measure of the apparent brightness of the test field would thus be influenced by this third variable of binocular contrast. Fry and Alpern (**4**) have demonstrated that monocular contrast effects decrease proportionately as the spatial separation between the inducing and the test field increases. Graham (**5**) gives a detailed account of the problem of interaction in many of its manifestations.

Method.—To test the possibility of binocular interaction in the present experiment, the match-field luminance (B_m) was studied as a function of the inducing-field luminance (B_i) for a test-field luminance level of .60 log ml., and for the two *Ss* used in Exp. I. The procedure used was different for each *S*, however. For A. D., the contrast function was studied in three sessions in each of which a different apparent binocular separation (constant during each session) was

TABLE 1

APPARENT BRIGHTNESS OF THE TEST FIELD (B_m) IN LOG MILLILAMBERTS AS A FUNCTION OF INDUCING-FIELD LUMINANCE (B_i), FOR DIFFERENT TEST-FIELD LUMINANCES (B_t)
(Data are for subjects H.L. and A.D.)

Test-Field Luminance (log B_t)																				
2.71			2.17			1.62			1.15			.60			.06			−.39		
Log B_i	Log B_m		Log B_i	Log B_m		Log B_i	Log B_m		Log B_i	Log B_m		Log B_i	Log B_m		Log B_i	Log B_m		Log B_i	Log B_m	
	H. L.	A. D.		H. L.	A. D.		H. L.	A. D.		H. L.	A. D.		H. L.	A. D.		H. L.	A. D.		H. L.	A. D.
−∞*	*2.63*	*2.80*	−∞	*2.19*	*2.23*	−∞	*1.56*	*1.67*	−∞	*1.16*	*1.21*	−∞	*0.61*	*0.65*	−∞	*0.03*	*0.14*	−∞		*−0.41*†
−1.59	2.72	2.69	−1.50	2.05	2.21	−1.45	1.53	1.64	−1.50	1.30	1.12	−1.51	0.65	0.58	−1.50	−0.05	0.20	−1.39		−0.45
−0.39	2.75	2.62	−0.48	2.13	2.23	*−0.82*		*1.64*	*−1.50*		*1.16*	*−0.94*		*0.60*†	*−1.50*		*0.12*†	−0.39		−0.46
−0.06	2.74	2.72	*−0.48*		*2.16*	−0.46	1.60	1.55	−0.41	1.28	1.21	−0.43	0.65	0.56	−0.48	0.01	0.17	*−0.39*		*−0.43*
−0.06		*2.64*	0.06	2.19	2.27	0.06	1.68	1.64	*−0.41*		*1.19*†	0.06	0.66	0.65	0.06	0.00	0.03	0.05		−0.60
0.60	2.71	2.69	0.61	2.19	2.29	*0.06*		*1.70*†	0.12	1.19	1.25	0.60	0.74	0.55	*0.06*		*0.13*	0.64		−0.81
1.15	2.70	2.77	*0.61*		*2.30*†	0.59	1.63	1.60	0.61	1.19	1.25	*0.60*		*0.54*	0.60	−0.37	−0.24	*0.64*		*−0.83*†
1.15		*2.62*†	1.14	2.16	2.24	1.08	1.63	1.58	1.15	1.17	1.24	1.14	0.31	0.40	1.09	−0.53	−0.30	1.17		−1.19
1.68	2.69	2.72	1.63	2.23	2.26	1.62	1.65	1.67	*1.15*		*1.20*	1.63	0.13	0.15	*1.09*		*−0.37*†	1.72		−1.25
2.17	2.66	2.58	2.17	2.09	2.25	*1.62*		*1.62*	1.69	0.73	0.73	*1.63*		*0.15*†	1.62		−0.52	2.26		−1.47
2.71	2.66	2.67	*2.17*		*2.20*	2.16	1.42	1.35	2.18	0.60	0.66	2.16	0.05	−0.04	2.17		−0.74			
2.71		*2.66*	2.71	1.67	1.82	2.65	1.12	1.28	*2.18*		*0.69*†	2.71	−0.15	−0.34						
						2.65		*1.18*†	2.71	0.18	0.40									

* Italicized figures are from check experiment.
† These represent means of data obtained in two sessions.

introduced between the match field and the test and the inducing fields. These separations were 43′, 70′, and 119′, and were made possible by a lateral (nasalward) displacement of the match-field fixation point from the match field. When the match-field fixation point was moved 98′ from the match field and then fused with the test-field fixation point, the match and test fields appeared to be separated by 119′ in the binocular view. For the displacement of 42′, the data used are those obtained in Exp. I at the test-field luminance of .60 log ml. Data for each separation were obtained in each session by a procedure similar to that of Exp. I for test-field luminances below 10 ml. The session involving 70′ separation was repeated and the data averaged with the data for the first 70′ separation.

For H. L., a different procedure was followed in which apparent binocular separation was the independent variable rather than the inducing-field luminance. That is, at the beginning of an experimental session *S* dark adapted 10 min., then light adapted to the test field (.60 log ml.) and lowest inducing-field luminance for that session. This initial adaptation was at the separation of 42′; at this separation, four practice and ten experimental matches were then made. Two minutes of dark adaptation followed, after which *S* was light adapted and made four practice and ten experimental matches for a 70′ separation. Another 2 min. of dark adaptation followed and the same procedure was followed for the 98′ separation. Thus, while the separation was increased in these initial three sets of ten matches, the inducing- and test-field luminances were held constant. The binocular separation was then readjusted to 42′, and the inducing-field luminance increased. The *S* then adapted and made matches, as before, for the three successive separations. During one experimental session three such sets of matches were made so that, in three sessions, apparent-brightness values were obtained for the three separations at each of nine values of the inducing-field luminances, starting from zero luminance at the beginning of the first session, to 2.71 log ml. at the end of the last session.

TABLE 2

APPARENT BRIGHTNESS OF THE TEST FIELD (B_m) IN LOG MILLILAMBERTS AS A FUNCTION OF INDUCING-FIELD LUMINANCE (B_i) FOR THREE APPARENT BINOCULAR SEPARATIONS BETWEEN THE MATCH AND THE TEST FIELDS

(The test-field luminance is held at .60 log ml.)

Inducing-Field Luminance (log B_i)	Separation in Minutes of Visual Angle					
	119′		70′		42′	
	H. L.	A. D.	H. L.	A. D.	H. L.	A. D.
$-\infty$	.47	.33	.61	.52	.75	.66
−1.51	.64	.32	.76	.49	.08	.58
−0.43	.44	.34	.64	.51	.68	.56
0.16	.40	.43	.64	.53	.76	.65
0.60	.47	.29	.65	.40	.64	.55
1.14	.02	.09	.15	.24	.33	.40
1.63	−.06	−.06	.06	.14	.08	.15
2.16	−.15	−.16	−.07	−.08	.05	−.04
2.71	−.54	−.45	−.40	−.43	−.19	−.34

FIG. 3. Test-field apparent brightness (B_m) as a function of inducing-field luminance (B_i) for three apparent binocular separations between the match and test fields. The test-field luminance (B_t) is .60 log millilamberts for each curve.

Results.—Both *S*s, therefore, made matches for the same inducing- and test-field values for each separation, but under different procedures. Their data are averaged and entered in Table 2 and graphed in Fig. 3. The curves in Fig. 3 are empirically fitted to the data.

The basic argument of the experiment is as follows: We know that interaction effects in one eye are a function of the distance separating the interacting stimuli; the greater the separation, the less the degree of interaction (**2, 4**). Presumably, too, the binocular interaction effects are a function of the retinal disparities that separate the stimulations in the two eyes from corresponding points: the greater the disparity, the greater the "apparent" separation and the less the interaction.

The curves in Fig. 3 show apparent brightness of the test field (B_m) as a function of the inducing-field luminance (B_i) for the three apparent separations between the match field and the test and inducing fields. The curves are identical in shape to that curve in Fig. 2 for which the test-field luminance is equal to .60 log ml. The entire curve is displaced downward, however, as the separation between the match field and the test and inducing fields increases. In other words, as the distance of the match field from the test and inducing fields is increased, the apparent brightness of the test field shows a slight decrease over the entire range of inducing-field luminances. The question arises: Are the curve displacements in Fig. 3 due to a change in binocular interaction between the inducing and the match fields as the latter is displaced laterally? In reply to this question, one point is worthy of emphasis.

The square symbols associated with each curve of Fig. 3 represent the match value of B_m (against B_t) in the absence of an inducing field. The relatively flat portion of each curve is at the level given by the match, and the shape of each curve is unchanged by separation. Since we cannot speak of interaction in the absence of the inducing field, it appears that, because the position of each curve with respect to the match made in the absence of such a field remains constant, we cannot speak of inducing-field and match-field interaction.

The successive downward shift of the function, however, remains to be explained. A first tentative hypothesis would attribute the shift to the fact that as the match field is displaced, it moves to a more sensitive retinal area. It is well known that retinal sensitivity increases in the periphery (**7**). The side of the match field nearest the test field is displaced in the present experiment through 98′ from the foveal "fused" fixation spot; actually the far side extends to 131′ from the fixation spot since the field is 33′ wide. It is at this maximum displacement that the function's downward shift in Fig. 3 is most noticeable. It seems possible, therefore, that at 98′ from the fovea the match field, extending to 131′, is stimulating a retinal area more sensitive than that stimulated by the test field. The match field would then appear slightly brighter than when at 21′ and the test field measured as dimmer along the entire contrast function. A final explanation of this effect, of course, would require a much more detailed investigation than is offered in this experiment.

Discussion

When we speak of simultaneous brightness contrast, what do we mean? Historically, it has involved the apparent brightness of one visual field relative to the luminance of an adjacent or surrounding field (**10**, p. 567–571). In the present experiment the test-field apparent brightness is thus related to the luminance of the adjacent inducing field.

There is more to consider in this relationship, however, than just these two fields. Since, if we speak of simultaneous contrast as the relation between one visual field and another, it becomes inconsistent to ignore the relation between either of these illuminated fields and the apparently black surround that is nevertheless active at the retinal level. The retinal activity of the surround might arise from entoptic scattering in the ocular media (**1**, p. 57–71) or from interaction effects in the retina as discussed by Graham (**5**).

If we, therefore, look to the activity taking place in the retina as a necessary event in an explanation of simultaneous brightness contrast, we must then consider the retinal activity in the apparently black, or nondirectly stimulated surround of the test field as well as that in the test- and inducing-field areas. And since the match field was matched in apparent brightness to the test field, its surround must also be considered.

Thus the following assumptions might be made fruitfully with respect to the apparent brightness of the test and match fields: (*a*) The apparent brightness of the match field is assumed to be proportional to the following ratio: the match-field luminance (B_m), raised to some power (a_1), divided by the activity (H_m) in the retinal surround. (*b*) The apparent brightness of the test field is assumed to be proportional to the following ratio: the test-field luminance (B_t), raised to some power (a_2), divided by the combined activity in the retinal surround resulting from stimulation of the test and inducing fields, or $H_t + H_i$. It should be noted that included in this combined surround activity ($H_t + H_i$) is that activity in the inducing-field itself.

On equating apparent brightnesses, therefore, of the test and match fields:

$$k\left(\frac{B_m{}^{a_1}}{H_m}\right) = k'\left(\frac{B_t{}^{a_2}}{H_t + H_i}\right). \qquad (1)$$

Assuming that for our experimental situation $k = k'$, these terms cancel out in equation (1). And also assuming that H is directly proportional to B then

$$H_m = k_1 B_m \qquad (2)$$

$$H_t = k_2 B_t \qquad (3)$$

$$H_i = k_2 B_i. \qquad (4)$$

Finally, changing H to its proportional B value, cancelling out and consolidating constants, and arranging the equation in logarithmic form we get

$$\log B_m = \frac{a_2 \log B_t - \log (k_2 B_t + k_2 B_i) + \log k_1}{a_1 - 1}. \qquad (5)$$

Equation (5) predicts the data in Table 1 and Fig. 2 satisfactorily when the values of the constants are as follows:[5]

$$a_1 = a_2 = 3.0$$
$$k_1 = k_2 = .0174$$
$$k_2 = .0064.$$

The above formulation has some merits and also some weaknesses. One of the merits, of course, is that the constants, just evaluated, have the same value for each curve in Fig. 2. Further, it seems capable of handling parameters additional to those of the inducing-field and test-field luminances.

For example, the slope of the function is proportional to the values of a_1 and a_2 in equation (5). These constants, therefore, might contain values representing the effect of the inducing-field area as a parameter of contrast, since in an exploratory experiment carried out by the writer, the slope of the contrast function was found to increase with an increase in the ratio of inducing- to test-field area. The effect of retinal separation between test and inducing fields would be calculated in $H_i[k_2 B_i$ in equation (5)], since H_i varies with retinal separation as well as inducing-field luminance, as seen in the study of Fry and Alpern (**4**). Whether or not variation of these parameters will produce data predicted by this equation is an empirical question. The point is, however, that in addition to predicting the data of the present experiment, the formulation does make testable provisions for relating these parameters under one system.

A weakness of the above formulation is that it tells us very little about the mechanism responsible for contrast. Something might be said, however, about the role of scattering as responsible for the effect in our curves in Fig. 2. As pointed out by Fry and Alpern, the contrast effect may be due to an inhibition of the test-field apparent brightness by the inducing-field scattered light, acting at the borders of the test field.

[5] For our experimental situation a_1 is assumed equal to a_2 and k_1 equal to k_2. The calculated values of these constants are arbitrary and can be estimated in the following manner: Equation (5) can be reduced to the general form

$$\log B_m = A - B \log (C + D B_i),$$

where

$$A = \frac{a_2 \log B_t}{a_1 - 1} - \frac{\log k_1}{a_1 - 1};$$

$$B = \frac{1}{a_1 - 1}; C = k_2 B_t; \text{ and } D = k_2.$$

This equation is a straight line in $\log B_m$ and $\log (C + DB_i)$. The values of C and D may be estimated such that the plot of $\log B_m$ and $\log (C + DB_i)$ yields a straight line. A is an intercept and B the slope in this equation.

Under such circumstances the effects H_m, H_t, and H_i, could be regarded as representing effects due entirely to scattering. For this to be true, however, the percentage of light scattered from the inducing field to the test field, according to the calculated value of k_3 in equation (4), would become relatively large for small values of test-field luminance. Thus, equation (5) would become:

$$\log B_m = \frac{a_3 \log (B_t + S_i) - \log[k_2 (B_t + S_i) + k_3 B_i] + \log k_1}{a_1 - 1}, \quad (6)$$

where H_m, H_t, and H_i now represent scattered light from the match, test, and inducing fields, respectively, and S_i, representing scattered light specifically falling on the test field, is equal to H_i.

It is of interest to note that the inclusion of the term, S_i, is unnecessary in order for our formulation to fit the data in Fig. 2. This suggests several alternatives to an hypothesis based upon light scattering in the optic media. First, and most probably, some sort of physiological interaction rather than scattering is at the basis of the contrast effect; the H terms in our equations are not determined entirely by the amount of scattering alone but by a combination of scattering and interaction, or by the latter alone. In this case, the value of S_i could be relatively small to H_i and also to B_i and thus negligible in the range investigated in our experiment; we could expect our curves to be of the form that they, in fact, are.

A second possibility is that the H terms in our equations actually do represent scattering, but that the subject is somehow discriminating the scattered light (S_i) added to the test field, and compensating for this added light in his brightness judgment. A discrimination of the scattered from the test light might be on the basis of its diffuseness, as different from the focused light of the test field. The compensation might be of the same sort described by Woodworth (10, p. 596): S in judging the brightness of a piece of coal under strong illumination, compensates for the added illumination and judges its brightness as much less than it actually reflects. The stimuli and responses for this "discrimination plus compensation" must, of course, be identified and explained in themselves. General information concerning this point, however, would seem of some use in determining whether or not contrast is primarily a function of entoptic scattering in the ocular media. If we assume, therefore, that the scattered light added to the test field is being discriminated and compensated for, the term, S_i, may be dropped from equation (6).

It should be pointed out that the foregoing arguments (against the effectiveness of scattered light in contrast) are developed from a theoretical formulation which itself may or may not be acceptable. Such acceptability awaits further investigation of the implications of this formulation.

Summary

Simultaneous brightness contrast is investigated in the fovea as a function of the inducing- and the test-field luminances. The inducing and the test fields are adjacent squares, each subtending 33′ on a side, and presented to S's right eye. A match field of equal size is presented to S's left eye. The S's task is to set the luminance of the match field so that its brightness appears equal to that of the test field. As the inducing field is varied in luminance from below to above that of the test field, the apparent brightness of the test field, and thus the contrast effect, is measured by the actual luminance of the match field.

The results show the contrast function to be of the following nature. The test-field apparent brightness is depressed only slightly for inducing-field luminances less than that of the test field. For inducing-field luminances greater than that of the test field the depression of the test-field apparent brightness seems directly proportional to the luminance of the inducing field. Thus, a "bend" in the function occurs when the inducing-field luminance is approximately equal to that of the test field. This relation is consistent over a large range of test-field luminances.

A theoretical formulation based upon the relation of the test-field apparent brightness to the activity in the visual field completely surrounding the test field leads to an equation satisfactorily fitting the data. The physiological nature of this surround activity is undetermined.

The possibility of binocular contrast effects between the inducing and the match fields was considered and shown to be insignificant in a control experiment.

(Received September 6, 1952)

References

1. Bartley, H. S. *Vision.* New York: Van Nostrand, 1941.
2. Beitel, R. J. Spatial summation of subliminal stimuli in the human retina. *J. gen. Psychol.*, 1934, **10**, 311–327.
3. Chapanis, A. How we see; a summary of basic principles. In *Human factors in undersea warfare.* Washington, D. C.: National Research Council, 1949.
4. Fry, G. A., & Alpern, M. The effect of veiling glare upon the apparent brightness of an object. *J. opt. Soc. Amer.*, in press.
5. Graham, C. H. Vision: III. Some neural correlates. In C. Murchison (Ed.), *A handbook of general experimental psychology.* Worcester: Clark Univer. Press, 1934. Pp. 829–879.
6. Guilford, J. P. *Psychometric methods.* New York: McGraw-Hill, 1936.
7. Hecht, S. Rods, cones, and the chemical basis of vision. *Physiol. Rev.*, 1937, **17**, 239–290.
8. Hess, C., & Pretori, H. Messende Untersuchungen über die Gesetzenässigkeit des simultanen Helligkeitskontrastes. *Graefes Arch. Ophth.*, 1894, **40**, 1–24.
9. Schouten, J. F., & Ornstein, L. S. Measurements on direct and indirect adaptation by means of a binocular method. *J. opt. Soc. Amer.*, 1939, **29**, 168–182.
10. Woodworth, R. S. *Experimental psychology.* New York: Holt, 1936.

Journal of Experimental Psychology
Vol. 45, No. 5, 1953

MOTOR PERFORMANCE AS A FUNCTION OF CLICK REINFORCEMENT[1]

BRADLEY REYNOLDS AND JACK A. ADAMS

Perceptual and Motor Skills Research Laboratory, Human Resources Research Center

In a study with the rotary pursuit task already reported (5), it was observed that *Ss* who were presented with a click for each period of 1-sec. continuous time on target performed significantly better than *Ss* who were not so treated. It seemed desirable to explore this effect further, and for this purpose a study was designed to investigate the effect of the length of the period of the continuous time-on-target period required for presentation of a click. It was not anticipated that any regular uniform relationship would be observed between duration of continuous on-target period required for click presentation and level of performance. Presumably, the duration of the continuous on-target period required could be made so brief that its effect would be negligible; or the duration could be extended sufficiently that the *S* would rarely, if ever, get on target for the continuous period required, and as a consequence would be presented with few, if any, clicks. It was also considered quite possible that there would be some interaction between the duration of the required continuous on-target period and the number of training trials. Presentation of a reinforcing click for a briefer continuous on-target period might conceivably have greater facilitative effect upon performance early in training than presentation of a click for a longer period. Early in training *S* is able to remain on target continuously for briefer periods but not for longer periods of time. With increased proficiency associated with continued training, however, *S* is able to stay on target for longer continuous periods of time. Thus, presenting clicks for short on-target periods might be effective early in training, while later a click for a longer on-target period would be more facilitative.

In designing the study, because of the expectations described above, an attempt was made to employ continuous on-target time intervals which it was hoped would include an optimum interval and intervals less than and greater than the optimum. Further, it was hoped that the longer and the shorter intervals employed would permit demonstration of the kind of interactions described above; that is, intervals were chosen which were felt to be short enough to lead to some facilitation early in training and long enough to lead to facilitation later in training. There seemed to be no good basis upon which an expected optimum interval could be established. It was felt that the optimum interval would not be as long as 2 sec. nor as short as .1 sec. for the number of trials employed in this study.

Procedure

Apparatus.—The apparatus was a rotary pursuit test of conventional design. It has been described previously (2). An electronic circuit was employed which could be made to activate a solenoid when *S* had accumulated a given increment of continuous on-target time. The continuous on-target time required for the activation of the solenoid could be varied. Activa-

[1] The experimental work for this study was performed as part of the United States Air Force Human Resources Research and Development Program. The opinions or conclusions contained in this report are those of the authors. They are not to be construed as reflecting the views or indorsement of the Department of the Air Force.

tion of the solenoid was accompanied by a loud distinct click clearly audible to *S*. Time on target in seconds was recorded on a Standard Electric Timer.

Subjects.—The *Ss* were 198 basic trainee airmen drawn from the population at Lackland Air Force Base. Thirty-three *Ss* were assigned to each of six groups.

Experimental groups.—Five experimental groups were differentiated on the basis of duration of continuous time on target required for presentation of a click. The following periods were employed: .10, .20, .50, 1.0, and 2.0 sec. Prior to differential experimental treatment all *Ss* were given five trials on the rotary pursuit test. These trials were 30 sec. in duration with 15 sec. between each trial. Total score for these five trials was determined for each *S* and, on the basis of the scores so obtained, *S* was assigned to one of 33 levels based upon initial rotary pursuit proficiency, and at random to one of five experimental groups. These 33 levels were established in a preliminary study on a multiple-unit rotary pursuit apparatus where 533 *Ss* were administered five 30-sec. trials with 15-sec. rest between trials. Cumulative five-trial scores were put into a frequency distribution and, on the basis of this distribution, 33 scoring intervals were established within which it was anticipated a high percentage of subsequent *Ss'* scores would fall.

Three minutes following the first five classification trials, all *Ss* were administered 50 30-sec. trials with 15-sec. rest between each trial. The *Ss* were told that they would hear a click whenever they remained on target continuously for the specified period of time. In the instructions for the five classification trials *Ss* had been told that their score depended upon the total amount of time they were on the moving target.

At the completion of the 50 acquisition trials all *Ss* were given five trials without presentation of click. Following the no-click trials, five additional trials were given where all *Ss* received a click for each .5-sec. continuous time on target. Appropriate instructions preceded the change in experimental conditions.

Control group.—For purposes of comparing the results obtained with the experimental groups, a control group was run using a multiple unit setup for rotary pursuit. These *Ss* were run immediately following the experimental groups. A total of 130 *Ss* were run through the total training sequence; that is, five pretest trials, a 3-min. interval, 50 acquisition trials, and then 10 trials, all without click. From these 130 *Ss*, 33 were selected who had cumulative scores for the first five trials falling in the 33 interval groups. When several *Ss* had scores falling in a particular interval, one *S* was chosen for this interval by the use of a table of random numbers. These *Ss* were given the same instructions as the experimental groups except for the reference to the click presentation.

Results

The precision of the matching technique can be observed in Table 1. It will be noted that the means and variances of the total score on the five classification trials for all groups are essentially equal.

The main results of the study are shown in Fig. 1. This figure shows mean time on target for all experimental groups for the last five acquisition trials and the five trials of each of the two post-acquisition test periods, scores for each five trials having been combined into a single mean value. Values for the control group on these three blocks of trials are indicated on the ordinate. Best perform-

TABLE 1

Means and Variances of Total Score on Five Classification Trials for All Groups

Group	Mean	σ^2
.1 sec.	28.78	275.97
.2 sec.	29.08	273.30
.5 sec.	28.78	268.54
1.00 sec.	29.01	268.34
2.00 sec.	29.23	274.41
Control	28.83	267.38

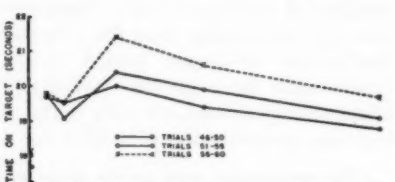

Fig. 1. Performance level for the last five acquisition trials (45–50), the five no-click trials (51–55), and the five trials where all experimental groups received a .5-sec. click (56–60)

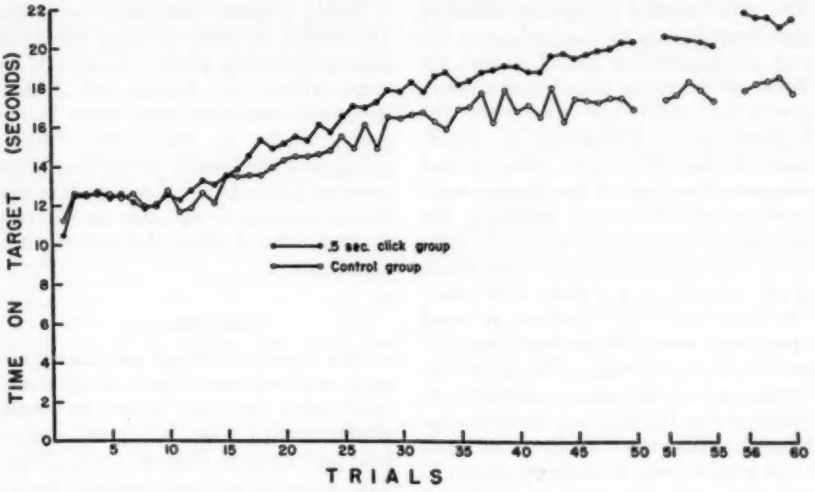

FIG. 2. Performance curves for the .5-sec. group and the control group. The scores for the first five classification trials are not represented.

ance throughout is shown by the .5-sec. group. Among the experimental groups the 2.0-sec. group performs at a level below that for the other experimental groups during Trials 1–50. However, during the two post-acquisition periods, the performance for the 2.0-sec. group is indistinguishable from that of the .2-sec. group. During the acquisition period the .10-, .20-, and 1.0-sec. groups do not show any great differences between themselves, but during the two post-acquisition periods the 1.0-sec. group performs at a level higher than all other experimental groups with the exception of the .5-sec. group.

Figure 2 shows the mean time on target for the .5-sec. group and the control group during the acquisition trials and the two post-acquisition periods. Once differences between the groups are established, the .5-sec. group is clearly superior to the control group at all stages of practice. During Trials 51–55 where the click is removed, the .5-sec. group shows a slight tendency to decrease in level of performance but there is little indication that this group would have stabilized at the level of the control group if practice without click had continued. The presence of the click during acquisition has apparently led to a stable superiority over the control group. During Trials 56–60 where all experimental groups practice with a .5-sec. click, the .5-sec. group displays a sharp increase in performance level when returned to the conditions of practice prevailing during the acquisition period.

The other experimental groups were also superior to the control group throughout but were below the performance level of the .5-sec. group. Three treatment-by-levels analyses of variance (1) were made using experimental groups only since the concern in this study was primarily with the effect of click reinforcement for various amounts of continuous on-target time.

The experimental groups are differing significantly among themselves at the end of acquisition (total scores for Trials 46–50 being taken as the measure of performance). The F for the between-group differences is significant at the 5% level. The second analysis of variance is based upon total score on the five trials following the acquisition series (Trials 51–55). The F for the between-groups differences is not significant for these five trials. The third analysis of variance is based upon total score for the final five-trial period (Trials 56–60). The F for the between-group differences is significant between the 1% and 5% level of confidence. The three analyses of variance would seem to indicate that the duration of the on-target period required for presentation of click contributes significantly to between-group differences, so long as a click is present in the situation. With the removal of the click the differences observed are not significant, although the .5-sec. group maintains its superiority.

Three analyses of variance were also made which included the control group. As might be anticipated, the analysis for the five trials at the end of acquisition, the five no-click trials, and the final five trials with .5-sec. interval revealed F ratios significant beyond the 1% level of confidence.

TABLE 2

MEAN SCORE FOR SELECTED BLOCKS OF FIVE TRIALS

Group	Trials		
	1–5	26–30	46–50
.1	11.15	17.27	19.83
.2	12.02	16.82	18.11
.5	12.12	17.53	20.46
1.0	11.87	16.74	19.44
2.0	11.79	16.12	19.17
Control	12.34	15.87	17.24

Table 2 shows the mean score for the groups on three selected blocks of five acquisition trials. Contrary to expectations, the .10-sec. and .20-sec. groups do not show superiority in the early trials. In fact, the .10-sec. group had the lowest level of performance on Trials 1–5. A treatments-by-levels analysis of variance on the two early blocks of trials did not yield a significant F.

DISCUSSION

The results obtained are consistent with expectations that an intermediate click interval would be most effective since .5 sec. was found to give best performance throughout training. The anticipation with respect to an interaction between interval length and number of trials was not borne out in the results. The optimum interval observed, that is .5 sec., produced best performance throughout most of the acquisition trials.

Since it was found that the performance of the click groups as compared with the control group was superior for the trials when the click was removed, it seems that the differences established by the presentation of the click for various continuous on-target intervals is associated to some extent with differences in habit strength. That some of the differences observed are motivational in character, however, is not inconsistent with the results.

Turning to an interpretation of the results, it becomes obvious that the results admit of several theoretical interpretations. Certain of these will be suggested in what follows. It is felt that insufficient data exist at the present time to permit any rigorous theoretical interpretation.

It is clear that the results might well be interpreted within the framework of Thorndikian reinforcement theory (6). The presentation of the click, in the present study, could be conceived as equivalent to the confirmatory event in the Thorndikian situation where E says "Right" or "Wrong" with the occurrence of a given response.

The results, of course, admit of a motivational interpretation since a goal for performance beyond the usual one of simply remaining on target as much of the time as possible is established through the requirement for continuous

on-target periods for click occurrence. As noted above, the persistence of the effect of click presentation during the period when no click is present would seem to indicate that, although motivational components may be present, there are some effects which seem more legitimately attributable to learning.

The results can also be interpreted in relationship to the characteristics of the particular continuous pursuit task employed, and the study was designed with this primarily in mind. In the earlier study involving presentation of click for continuous on-target time, an interval of 1 sec. was employed. The choice of this interval was dictated by consideration of the fact that the rotary pursuit task requires *S* to make a regular smooth circular movement at the rate of one per second in order to remain on target; and although the rotary pursuit task is a continuous one, it can be conceived as consisting of a succession of smooth rotary movements at a 1-sec. rate. The results obtained can be interpreted as being consistent with this assumption. The shorter continuous on-target intervals, that is, .1 and .2 sec., are of insufficient duration to force upon *S* a continuous circular movement. The .5-sec. interval, however, does require that *S* make a considerable part of the rotary movement required. It may well be that only a portion of the rotary movement required by the rotary pursuit task can be acquired by *S* under conditions holding in the present study, whereas the more exact rotary movement required by the 1- and 2-sec. intervals cannot. Since there is presentation of a click for a given period of continuous on-target time, the results can be interpreted within the framework of knowledge of results. The experiment, so conceived, can be considered to be one in which additional knowledge of results, a click, is presented for responses varying in difficulty or precision as defined by continuous on-target time. The results would indicate that such knowledge of results produces improved performance and has its maximum effect when the response is of some intermediate level of difficulty.

An interesting parallel exists between the results of this study and the results of certain conditioning studies. It has been observed in a number of conditioning experiments (**3, 4, 7**) that the optimum interval between the occurrence of a conditioned stimulus and the presentation of the unconditioned stimulus is of the order of .5 sec. As the interval is shortened or lengthened with respect to .5 sec., the strength of conditioning as measured by, for example, frequency of response, is decreased. In the present study the reinforcing click emphasizing the on-target event has its maximum effect when the interval separating the occurrence of the click and the response getting *S* on target is .5 sec.

No attempt will be made at the present time to analyze the rotary pursuit task in terms of conditioning. Even though the interval yielding optimum performance in this study is approximately the same as that found in conditioning, it is possible that this correspondence may be fortuitous. The similarity, however, between the functions shown in Fig. 1 and those showing the relationship between performance and the length of the CS–US interval in conditioning is sufficiently close to warrant serious consideration.

Summary

A previous study on the rotary pursuit task (**5**) revealed that the presentation of a click reinforcement for continuous 1-sec. time-on-target period led to consistent superiority when compared with a control group not receiving the click. The present study investigated the effect of different continuous on-target times required for the presentation of click reinforcement.

The rotary pursuit task was employed. Five experimental groups were differentiated on the basis of duration of continuous time on target required for presentation of a click. Intervals of .10, .20, .50, 1.0, and 2.0 sec. were used. The *S*s were stratified on the basis of total score for the first five trials. These stratification trials were followed by 50 training trials in which click was delivered, 5 trials where click was absent, and 5 trials where all experimental groups received a click for each .5-sec. continuous time on target. A control group received the same number of trials without receiving click reinforcement at any time.

All experimental groups were superior to the control group at all stages of practice with the .5-sec. group generally displaying the highest level of performance throughout. The failure of the experimental groups to shift to the level of the control group during the no-click period suggests that the

superiority of the experimental groups can, in part, be attributed to learning rather than motivational variables.

Several theoretical interpretations were considered.

(Received September 8, 1952)

References

1. Lindquist, E. F. *Statistical analysis in educational research.* Boston: Houghton Mifflin, 1940.
2. Melton, A. W. (Ed.) *Apparatus tests.* AAF Aviation Psychology Program Research Report No. 4, 1947, 329–399.
3. Reynolds, B. The acquisition of a trace conditioned response as a function of the magnitude of the stimulus trace. *J. exp. Psychol.*, 1945, **35**, 15–30.
4. Reynolds, B. Extinction of trace conditioned responses as a function of the spacing of trials during the acquisition and extinction series. *J. exp. Psychol.*, 1945, **35**, 81–95.
5. Reynolds, B. Motor performance as a function of reinforcement and spacing of trials. *Amer. Psychologist*, 1951, **6**, p. 289. (Abstract.)
6. Thorndike, E. L. *Selected writings from a connectionist's psychology.* New York: Appleton-Century-Crofts, 1949.
7. Wolfle, H. M. Conditioning as a function of the interval between the conditioned and the original stimulus. *J. gen. Psychol.*, 1932, **7**, 80–103.

Journal of Experimental Psychology
Vol. 45, No. 5, 1953

ACQUISITION AND EXTINCTION OF AN INSTRUMENTAL RESPONSE SEQUENCE IN THE TOKEN-REWARD SITUATION[1]

WILLIAM W. LAMBERT, ELISABETH C. LAMBERT

Cornell University

AND PETER D. WATSON

Harvard Medical School

Studies on chaining of responses have showed the familiar goal-gradient phenomenon. The strength of S–R bonds in a series is presumed to be greatest close to the point of reinforcement, and the chain is presumed to be less strong close to the beginning of the response series. This differential strength of conditioning has been considered in a paper by Lambert and Solomon (6), wherein the extinction of a running response in rats was studied as a function of distance of block point from the goal. The results showed increased resistance to experimental extinction as the point of extinction approached in space and time the original point of reinforcement.

The present paper is concerned with formally similar studies with children, utilizing a token-reward situation. The children's behavior in the acquisition, as well as the extinction, situation has been studied.

Two sets of experiments will be described. The first will be noted more briefly, since it is felt that a specific uncontrolled, but interesting, variable in this situation confused the results obtainable. The second experimental situation was an outgrowth from the findings, and the deficiencies, of the first.

Apparatus

The Solomon token-reward machine was used. A diagram of this apparatus is given in Fig. 1. The *S*s turned the handle on the machine a specific number of turns; received a token, that is, a poker chip, from the token chute; inserted the token in the insert slot; and, after varying numbers of these act sequences, received a small piece of candy from the candy trough. The

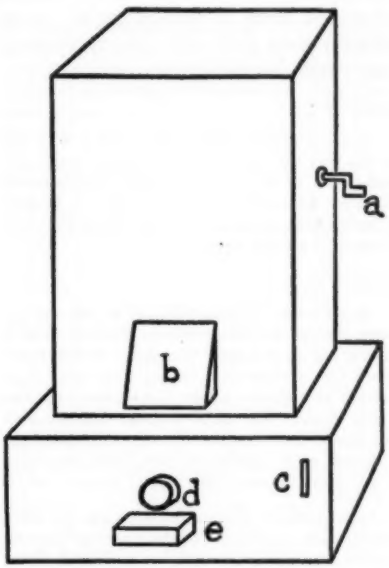

Fig. 1. The token-reward vendor. *a*—turning handle; *b*—token chute cover; *c*—token insert slot; *d*—candy chute opening; *e*—candy trough

[1] This work was done under the general direction of Dr. R. L. Solomon and was facilitated by the Laboratory of Social Relations, Harvard University. The data are included as a portion of the experimental work reported in the senior author's Ph.D. thesis (1950). The authors appreciate suggestions on exposition from Dr. Harold Schlosberg, Brown University.

TABLE 1

OUTLINE OF EXPERIMENT I

Group	Crank →	Click$_1$ →	Token →	Insert →	Click$_2$ →	Candy
Extinc.						
A	X	BP*—	—	—	—	—
B	X	X	BP—	—	—	—
C	X	X	X	X	BP—	—
D	X	X	X	X	X	BP—
Control						
X	X	BP (not used in training)			—	—
Y	X	(not used in training)				BP—

* BP = Block point.

entire sequence was repeated specific numbers of times, as will be described below.

EXPERIMENT I

Purpose

The aim of this study was to examine the results of experimental extinction on a chain of instrumental responses performed by children. Specifically, it was to study the results obtained from a procedure of using block points different distances along the chain from the goal.

Subjects[2]

Two groups of children were used. The first group was composed of 37 children, between 3 and 5 yr. of age, from the Harvard Nursery School. The second was composed of 24 children from the Salvation Army Nursery School in Boston, of similar ages.

Procedure

Acquisition.—The basic plan for this experiment is given in outline form in Table 1. The *S* turned the work handle 18 times. At this point there was a clear click sound from the token vendor and then a token was ejected into the token chute. The child inserted the token in the insert slot. Another clear, but distinctively different, sound came, this time from the candy vendor, and following this a small piece of candy was automatically ejected into the candy trough. The experimental *S*s in the first group of children went through this procedure once a day for ten days. The experimental *S*s in the second group went through it five times a day for ten days. For the control *S*s in both groups, the procedure was similar to that of the experimental *S*s, except that no token was used in the sequence to candy. The handle on the token vendor was turned 18 times by these control *S*s, and with accompanying noises of the candy vendor (Click 2 in Table 1), the candy was automatically ejected.

It should be noted here that in this experiment, and also in the second study cited below, the group design is in some respects more complex than necessary for obtaining the data here reported. The reason for this fact is that in each case the difficulty in securing sufficient *S*s of similar ages prevented some of the planned procedures from being carried out as anticipated.

Extinction.—Four extinction groups were obtained from the experimental *S*s. Group A was composed of children who turned the work handle but who were rewarded by neither the receipt of the token or even the usual click of the work machine. They could continue to turn the handle, but no result ensued. Group B was composed of children who turned 18 turns, heard the click, but received no token. They could continue to turn and hear the click. Group C *S*s turned, heard the click, received the token, inserted the token, but received no candy and were not permitted to hear the usual noise of the candy vendor. They could continue to repeat this sequence. Group D *S*s went through the whole situation as in the learning phase of the experiment, the only difference being that they received no candy reward. Their sequence could also be repeated.

Two control extinction groups were used. Group X turned the work-machine handle but did not hear the usual noise of the candy machine

[2] The authors wish to acknowledge the kind cooperation of the staffs of the Harvard Nursery School, and the Cambridge Lower School, Cambridge; of the Salvation Army Nursery School, Boston; of the Laboure Nursery School, South Boston; and of the Sewall Avenue Nursery School, Brookline, Mass.

and did not receive candy. Group Y heard the noise of the candy machine but received no candy.

Results

Table 2 presents the results of the procedure outlined above, giving medians and ranges for units of 18 turns for all experimental groups.

The most clear-cut finding is that Groups A and B resisted extinction longer than did Groups C and D, which were closer to the point of original reinforcement. This result is highly significant according to Wilcoxon's sum tests (7). It may be noted also that Groups A and B were blocked from proceeding further in the instrumental chain than turning the work handle, while Groups C and D were blocked from the reward but were permitted to insert the token.

Other differences, although interesting in their bearing on the action of secondary reinforcers, are merely suggestive and will not be discussed in detail here.

Conclusions

Before any implications for theory were drawn from these data, the *E*s felt it wise to redesign the experiment with an eye to controlling a variable, here not controlled, which might have crucially influenced the results. This variable was the kind of act involved at the block point.

TABLE 2

FREQUENCY OF RESPONSE (UNITS OF 18 HANDLE TURNS) AS A FUNCTION OF VARIATIONS IN BLOCKING POINT

Extinction Group	First Subgroup (N = 37)		Second Subgroup (N = 24)	
	Median	Range	Median	Range
A	14.0	8.5–20.3	34.6	10.3–50.0
B	18.2	3.0–37.0	17.0	16.0–21.0
C	1.0	none	1.0	1.0–30.0
D	1.0	1.0–2.0	5.5	1.0–19.0
X	21.3	17.5–45.0	15.0	7.0–38.0
Y	9.0	1.0–31.0	19.7	3.3–37.0

EXPERIMENT II

Purpose

The aim here was to investigate again the problem stated for the first experiment and also to consider more thoroughly the acquisition as well as the extinction phenomena.

Subjects

Thirty-one children, of ages 3 to 5 yr., were utilized. These *S*s were obtained through the nursery schools they attended, three schools in or near Cambridge, Massachusetts. Extinction groups were equated as nearly as possible according to age, sex, and nursery school group.

Procedure

Acquisition.—Table 3 gives a schematic outline of the plan of the experiment. Acquisition procedure can be seen by reading the top line of the outline. The *S* turned nine turns and this turning was followed by a click which accompanied the automatic ejection of a red token. The *S* then inserted the red token in the slot of the candy vending machine, and then turned the

TABLE 3

OUTLINE OF EXPERIMENT II

Extinc. Groups	Crank	→	Red Token	→ Insert	→ Crank	→	White Token	→ Insert	→ Candy
W	X		BP*—	—	—		—	—	—
X	X		X	X	BP—		—	—	—
Y	X		X	X	X		BP—	—	—
Z	X		X	X	X		X	X	BP—

* BP = Block point.

handle of the work machine nine more times. This time a white token was automatically ejected, and insertion of this in the token slot was followed immediately by ejection into the candy trough of a small piece of candy.

The number of such sequences per day was varied from one to four. The experiment lasted for ten days, with a total of 28 reinforcements. At the end of each child's trials for a given day, a blue token was used to end the sequence. This blue token was automatically ejected following a turning sequence, as were the other tokens. With one group of children the blue token followed a red token and with another group it followed a white token, with the one exception that on the first day of the experiment the blue token followed a white for all children. This latter procedure was used in order to have the initial presentation sequence, with the *E*, the same for all *S*s. The reason for the inclusion of the blue token in this experiment was that it was to have been part of a more elaborate design than was possible to carry out, because of the difficulty of obtaining *S*s.

Turning-time measures were recorded throughout. This was done by a second *E*, who sat in an adjoining room and read the turnings from a dial which led by a flexible cable to the work machine in the experimental room.

Extinction.—On the eleventh day of the experiment the *S*s were divided into groups for extinction purposes, as is shown in Table 3. Group W was extinguished at the beginning of the sequence. Work done prior to the extinction criterion of 2 min. of no handle turning was recorded in terms of units of nine turns each. Group X was blocked at the point in the sequence where the red token was inserted; that is, no matter how many sequences of nine turnings of the work handle the child went through, he received each time a red token and never achieved the white token. Group Y turned for and received a red token, but no amount of further turning resulted in receiving the usual white token, or even another red token. Group Z received the red token after the usual turning and then turned and received a white token, but without the insertion of this being followed by the usual candy reward. Further work behavior of the *S* resulted in repetitions of this sequence.

It may be noted that in the cases of the two groups which were extinguished while continuing to receive tokens, the number of trials was necessarily limited by the token capacity of the token vendor, which was 60 chips.

The measure of resistance to extinction recorded was the number of turning units of nine turns each, until an extinction criterion of 2 min. of no turning was reached.

With this amended procedure, comparisons which could not be made in the first experiment were now possible. The group blocked at the turning point nearest the start could be compared with a group blocked at a turning point nearer the goal; the group blocked at insertion near the goal could be compared with a group blocked at insertion at an earlier point. Thus it was hoped to control more nearly the variable of kind of act performed at block point.

Results

Turning-time measures.—In plotting the turning-time measures, it was found that the usual skewed distribution was present. Certain of the numbers were converted into logs for statistical analysis. The turning times for the first red and the first white tokens for each child each day were analyzed. Table 4 shows the results of analysis of the $\log_{10}$ of these data for the last three days, giving a *t* test of the significance of the mean difference between the turning for the red, versus the white token, and the .05 and .01 levels criteria. The *t*'s for the eighth and ninth days are significant at the .05 and .01 levels, respectively. The *t* for Day 10 is significant only at the .1 level.

Table 5 gives the results of *t* tests of the data for the last (tenth) acquisition day as compared with the second, which was the *S*'s first day without assistance from *E*. These tests show evidence of a learning effect, although the evidence is more clear-

TABLE 4

t TESTS OF SIGNIFICANCE OF DIFFERENCE BETWEEN LOG_{10} TURNING TIMES—DAYS 8, 9, 10, EXP. II

	Red(M) minus White(M)	*t**
Day 8	.0654	2.076
Day 9	.0823	2.928
Day 10	.0575	1.748

* t = 2.042 and 2.750 for the .05 and .01 levels of confidence, respectively.

cut with the turning times for the red token.

These acquisition data thus seem to be in the direction of the predictions of Hull and others with respect to experimentation with rats, that response strengths in a chain increase in intensity as the goal is approached. This can be considered as at least analogous to Hull's (**4**) and others' findings of a gradient of locomotion, as well as Arnold's (**1**) finding of a decreased reaction time in a chain of acts as the goal is approached. In addition, there is evidence that the speed of turning increases in somewhat of a learning function over a sequence of trials.

Measures of resistance to extinction.—The results of the measure of resistance to experimental extinction are shown in Table 6. There is no general trend for all four extinction groups, rather those groups extinguished while turning seem to do more work before extinction is complete than do the groups blocked in the chain after inserting a token. These differences are in the direction shown by the first experiment. These data, however, are not here statistically significant.

The important result is that the two groups which were blocked close to the goal resisted extinction more than did those comparable groups blocked farther from the goal. This difference is significant to beyond the .05 level of confidence by use of Festinger's test (**3**). Thus it appears that the results here show the same phenomenon as shown in experiments with lower organisms, namely, that other things being equal, there is increasing resistance to extinction as the extinction point approaches the point of original reinforcement.

TABLE 5

t Tests of Significance of Difference between Log_{10} Turning Times—White and Red Tokens

Token	Day 2(M) minus Day 10(M)	t*
Red	.1003	2.289
White	.0801	1.876

*t = 2.042 and 2.750 for the .05 and .01 levels of confidence, respectively.

TABLE 6

Ranges and Medians for All Groups: Exp. II

Measure	Extinction Groups			
	W	X	Y	Z
Range	15–50	3–52	13–60	2–60
Median	25.0	10.5	45.0	37.0

Discussion

It might appear that these findings, as well as the similar findings in a study on rats in a formally similar situation (**6**), are easily subsumable under Hull's model (**5**) as a case of the goal-gradient phenomenon. This is not the case, as has been pointed out with reference to the rat data on this question. The present Hull model is not clear in its predictions in such a situation. The goal-gradient aspect of the model would predict an increase in resistance to extinction as the block point approaches the goal. The inhibitory aspect of the model would predict the opposite, however, since those *S*s which have done the most work should extinguish the more quickly. When these two aspects are put together, the prediction might be (*a*) that there should be no difference in the work done by any of the extinction groups, or (*b*) that the group which is extinguished most close to the goal should resist extinction longer. The operational specificity of Hull's model does not permit working out this problem quantitatively for our data.

It may be possible that some addition to Hull's model along the lines suggested by Brown and Farber (**2**) and others may be necessary to generate these data and similar data obtained from rats. This addition would take the form of a postulation of increased "frustration drive" as the block point approaches the goal. This greater drive may function to hold off the occur-

rence of a resting response and consequently to increase the work done before extinction. The postulation of such a drive would generate the present data as well as the somewhat comparable data obtained from rats.

The problems that these data outline for Tolman's model have been outlined elsewhere (5) and are similar to those of the Hull model.

Summary

Two experiments are reported which test with children a finding reported earlier with rats: that as the extinction point approaches the goal, there is increased resistance to experimental extinction. A token-reward situation was utilized in both experiments. In the first experiment the *Ss* received a white token after turning a work crank. In the second experiment the sequence was lengthened so as to involve two different colored tokens, the second of which only was exchanged for reward. In both experiments *Ss* were divided into groups which were extinguished at different points along the two sequences of behavior. The results were as follows:

1. Where the sequence contained only one token, those *Ss* extinguished close to the goal, while inserting tokens, resisted extinction less than did those extinguished farther from the goal, while turning the crank. This finding was interpreted as showing possibly the effect of the type of act at the point of blocking. The second experiment was set up to control this factor.

2. When the sequence was expanded in Exp. II, more work was done before extinction by *Ss* who were extinguished close to the goal than by those extinguished farther from the goal. This effect is clear only when the type of act at point of blockage is controlled.

3. There is evidence for a learning function in the speed of turning for the token and for a discrimination between the two tokens in the longer sequence—the reinforced token was worked for with more vigor than was the token which did not lead directly to reward.

The extinction findings are discussed in terms of an hypothesis which emphasizes the importance of increased motivation strength due to frustration, which tends to delay the reinforcement of a resting response and thus delays extinction of the instrumental act sequence.

(Received August 11, 1952)

References

1. Arnold, W. J. Simple reaction chains and their interaction: I. Homogeneous chaining with terminal reinforcement. *J. comp. physiol. Psychol.*, 1947, **40**, 349–363.
2. Brown, J. S., & Farber, I. E. Emotions conceptualized as intervening variables—with suggestions toward a theory of frustration. *Psychol. Bull.*, 1951, **48**, 465–495.
3. Festinger, L. The significance of differences between means without reference to the frequency distribution function. *Psychometrika*, 1946, **11**, 97–105.
4. Hull, C. L. *Principles of behavior.* New York: D. Appleton-Century, 1943.
5. Hull, C. L. Behavior postulates and corollaries—1949. *Psychol. Rev.*, 1950, **57**, 173–180.
6. Lambert, W. W., & Solomon, R. L. Extinction of a running response as a function of distance of block point from the goal. *J. comp. physiol. Psychol.*, 1952, **45**, 269–279.
7. Wilcoxon, F. *Some rapid approximate statistical procedures.* New York: American Cyanimid Co., 1949.

Journal of Experimental Psychology
Vol. 45, No. 5, 1953

THE EFFECTS OF VERBAL AND NONVERBAL RESPONSES IN MEDIATING AN INSTRUMENTAL ACT[1]

WENDELL E. JEFFREY

Iowa Child Welfare Research Station, State University of Iowa[2]

An important aspect of human learning is the acquisition of language and the consequent effect of language on behavior. The clearest analysis of this has been made by Dollard and Miller who have stressed the importance of language in supplying human beings with cue-producing responses which function to produce a cue that is part of a stimulus pattern leading to another response. Thus they state that: "Attaching the same cue-producing response to two distinctive stimulus objects gives them a certain *learned equivalence* increasing the extent to which instrumental and emotional responses will generalize from one to the other. . . . Conversely, attaching distinctive cue-producing responses to similar stimulus objects tends to increase their distinctiveness" (**3**, p. 101).

Similarly, Hull in discussing secondary generalization concludes that, ". . . the common-sense notion of similarity and difference is based upon the presence or absence of primary generalization gradients, whereas so-called logical or abstract similarities and differences arise from secondary, learned, or mediated similarities and differences, particularly those mediated by verbal reactions" (**6**, p. 194).

If one assumes that the conditioning between an internal stimulus and an overt response can take place in the same way as conditioning involving external stimuli, it is possible that the conditioning of a verbal response to an external stimulus can take place independently of the conditioning of a motor response to the internal stimuli produced by a verbal response. Thus, an external stimulus which hitherto has not been associated with a particular motor response, may now elicit the response through the mediation of the label.

That discriminations can be learned more quickly when labels (verbal responses) are attached to the stimulus objects has been indicated by Pyles (**10**), Gagné and Baker (**5**), and Rossman and Goss (**11**). Kuenne (**7**) and Alberts and Ehrenfreund (**1**) have demonstrated that transposition of response on a size-discrimination problem is more likely to occur when the children have in their verbal repertoire a relational statement with respect to size, and a study by Birge (**2**) indicated that if two different objects are assigned the same name by *E*, these objects are more likely to be responded to in the same manner by *S* than if the objects are given different names.

Although investigations of the mediational process have usually involved verbal responses, it is conceivable that any other response which produces sufficiently distinct proprioceptive stimulation might also serve in this manner. Most relevant to this hypothesis is the research of Shipley (**12**), and Lumsdaine (**9**), and the aforementioned study of Birge which demon-

[1] This paper is based on a dissertation submitted in partial fulfillment of the requirements for the Ph.D. degree in the Department of Child Welfare of the State University of Iowa. The writer is indebted to Professor Vincent Nowlis for advice and assistance throughout the investigation.

[2] Now at Barnard College, Columbia University.

strated that a new stimulus object, to which reaching responses were reinforced, was more likely to be called by the name that was attached to the previous object to which reaching responses had been reinforced.

The present research was designed to further evaluate the mediational process in a situation where *E* could train the mediational responses and to a great extent control the subsequent use of the responses in a discrimination problem. Three groups of *S*s were used, one provided with verbal mediational responses, and two, one of which was a control, provided with manipulative mediational responses. By mechanically obstructing the mediational response in the control group during the training and test series following its acquisition, it was possible to evaluate the importance of the training sequence associating the internal stimuli of the mediational response with the response to be mediated.[3] Comparison between the two experimental groups gave an estimate of the effectiveness of a manipulative response in a mediational capacity.

Children were employed as *S*s in order to minimize the number of mediating responses, sets, or associations that might tend to operate in an experimental situation to mediate, or otherwise cause transfer of response from one stimulus to another. Furthermore, it was assumed that of human *S*s, young children would be capable of the least amount of spontaneous verbalization with respect to the mode of solution, or the relationship between the several tasks involved in the experimental procedure.

[3] The possibility that covert muscular responses might still occur even though the gross movement was not permitted has not been overlooked. However, such responses were not thought to be as important a source of stimulation or as likely to occur as subvocal responses after verbal training.

Method

Subjects.—The *S*s were 56 children from the State University of Iowa Preschool Laboratories. They ranged in chronological age from 3 yr. 3 mo. to 5 yr. 6 mo. and in mental age from 3 yr. 3 mo. to 7 yr. 10 mo. Eight of these 56 children were dropped from the experiment after the initial training session because of failure to learn and/or rejection of the experimental situation.

The 48 remaining *S*s were arbitrarily divided into two MA groups of 24 each on the basis of an MA above or below 5 yr. 5½ mo. Within each MA group *S*s were randomly assigned to the experimental and control groups as well as to the order in which they would be tested.

Apparatus.—The stimuli consisted of three different shades of Hering gray papers, No. 3, 16, and 50. Number 3 was not quite white, No. 16 was gray, and No. 50 was black. The reflected light of these three stimuli as measured in apparent ft.-candles by a Macbeth illuminometer was 7.70, 2.47, and .35, respectively.

The stimulus exposure device was a Hull-type memory drum, the stimuli being affixed to the tape so that on each exposure a color patch of ¾ by 1½ in. was displayed. Three stimulus series were used and were affixed to a single tape so that each series could be selected by moving the aperture across the front of the drum. The three series were as follows: (*a*) random presentation of black and white stimulus patches; (*b*) random presentation of black and white, with a gray stimulus patch interspersed as every fifth stimulus; (*c*) random presentation of white, gray, and black.

The memory drum was in a circuit so wired that by pressing a push button switch the drum would operate through an open-close shutter cycle and stop. This allowed *E* to prepare *S* before each stimulus presentation.

The response mechanism consisted of a piece of iron rod ⅜ in. in diameter and 4 ft. long. One end of this was bent to form a handle which projected from the apparatus on the right side of the memory drum. This handle could be slid in and out with a total excursion of 12 in., and could be turned 45° to the left or right. It was possible to control the movement of the handle by placing differently grooved boards in the base of the support through which the rod slid. A smaller rod, projecting perpendicularly 1¼ in. from the handle rod, ran in these grooves and by changing the boards the handle could be pushed or pulled only, turned only, or set so that it had to be turned before being pushed or pulled.

A Standard Electric Timer was wired in the circuit so that the timer started when the shutter opened, and stopped, either when the handle was

moved 1 in. forward or backward from the center position, or when the shutter closed.

The reward delivery mechanism was electrically operated but manually controlled by *E*. Marbles were used as secondary reinforcement and were delivered into a box to the left of the memory drum. The *S* took each marble and placed it in a board containing 38 circular recesses, each holding a single marble.

Procedure.—The *S*s served individually in two sessions of approximately 25 min. each which were run on consecutive days whenever possible. However, owing to the intervention of weekends and absences the between-test interval varied from one to five days. This variation would not be expected to make any difference in the results since there was a retraining session on the second day. Session 1 consisted of initial training and initial test; Session 2 consisted of mediational training, retraining, and retest.

For the *initial training* in Session 1 white and black patches were presented serially in random order, and the reinforced response was the pushing or pulling of the handle. Direction of movement was balanced, half of the *S*s being reinforced for pulling the handle to the white stimulus, the other half being reinforced for pushing the handle to the white stimulus. The opposite response to the black stimulus was reinforced in each case. The *S* was seated in front of the memory drum with his right hand on the handle. After instructions, training proceeded to a criterion of 25 correct responses in 26 trials. Each correct response was reinforced immediately with a marble. When 38 marbles had been collected, they were cashed in for a cellophane-wrapped after-dinner mint. In case *S* earned more than 38 marbles before reaching the training criterion, he was given the piece of candy and asked to continue playing for another piece. In most cases, *S*s met the criterion within 25 to 45 trials.

Initial test data were obtained after reaching the above criterion by shifting the stimulus exposure aperture to the second stimulus series of black and white patches with a gray stimulus interspersed every fifth trial. In this series there was a total of 30 trials, the second, third, and fifth stimulus in each block of five being arbitrarily designated test stimuli. This provided six test responses to each of the three different stimuli. By starting at two different places in the series the order in which the test stimuli occurred, and their relation to previous stimuli in the order. was balanced. To eliminate the reinforcement or nonreinforcement of any particular response during the test series, the reinforcement procedure was changed. The *S*s were now told that no more marbles would be delivered until they were through but that they were to keep on responding as they had before and the experimenter would keep account of the number of marbles they earned. After 30 trials enough marbles were dispensed to obtain the candy regardless of *S*'s performance.

TABLE 1

REINFORCED RESPONSES DURING MEDIATIONAL TRAINING FOR THE VARIOUS GROUPS AND SUBGROUPS

Stimulus	Subgroups of Group I		Subgroups of Groups II & III	
	A	B	A	B
White	"White"	"White"	Left turn	Left turn
Gray	"White"	"Black"	Left turn	Right turn
Black	"Black"	"Black"	Right turn	Right turn

In Session 2 the stimuli for *mediational training* were white, black, and gray patches presented with equal frequency. The *S*s within each MA group had been assigned at random to three treatment groups (I, II, and III), each further subdivided into two conditions (A and B). Group I was trained to make the verbal mediational response "black" or "white" and Groups II and III learned the motor response of turning the handle left or right. The two subgroups (A and B) under each of the three treatments differed with respect to the response made to the gray stimulus. The *S*s in Subgroup A were reinforced for making the same response to gray as to white, while *S*s in Subgroup B were reinforced for making the same response to gray as to black (cf. Table 1). The direction of turning was counterbalanced in each of the motor groups. Training proceeded to a criterion of 25 correct responses in 26 trials.

During *retraining* the first stimulus series of white and black patches was utilized again in order to connect the stimuli of the mediational response to the original push-pull response while relearning the latter. This time, Group I was required to make the verbal mediational response prior to the push-pull response and Group II the turning mediational response before the push-pull response. Group III was used as a control for the effects of combining the mediational and original responses and was not permitted to make the turning response while relearning the push-pull response. All *S*s were trained to a criterion of 15 correct responses in 16 trials and were given a piece of candy at the end of the series regardless of the number of marbles earned. If any *S* received 38 marbles before reaching the criterion, he was given the candy and asked to continue playing for another piece.

In the *retest*, the *S*s were tested on the stimulus

series of white and black patches with gray on every fifth trial as in the original test. The effectiveness of retraining with the mediational response was measured here since Group I was instructed to continue naming the stimuli before moving the handle and Group II had to make the turning response before the push-pull response was possible. Otherwise the *Ss* were told to keep on responding as they had been and the individual trials were again not reinforced.

For all training series, *instructions* included (*a*) a demonstration by *E* of the response to be learned, and (*b*) direction and correction by *E* of *S*'s trial responses to each stimulus in the series being used. It is to be emphasized that neither the stimuli nor the direction of movement were ever named except to say, "When it is this color, move the handle this way, and when it is this color, move the handle this way." This statement was accompanied by the demonstration of the appropriate movement of the handle.

Any verbalizations or questions by *S* about the stimuli or direction of movement were ignored or interrupted by giving the ready signal for the next trial. In the few cases where *S* did not appear to be learning, another demonstration was given accentuating the relationship between the stimuli and the direction of movement.

Response measures.—The latency, frequency, and direction of response were obtained on 6 trials to the white stimuli, 6 trials to the gray stimuli, and 6 trials to the black stimuli in both the test and retest series each of which consisted of 30 trials.

Because of difficulties presented by unlimited latencies, and the fact that the retest latencies were inflated by the inclusion of the mediational response with the push-pull response, the latency measures were not analyzed. The frequency measures, however, were converted to a difference score reflecting the change in response to the gray stimuli from the test to retest trials. This score was obtained by arbitrarily assigning the value +1 to correct responses, i.e., responses which were in the same direction as mediational training for that *S*, and −1 to responses in the opposite direction. If no response was elicited within approximately 3 sec. after the shutter opened, a zero score was given. The sum of the values for the six trials was then obtained for the test and retest sessions separately and a difference score was computed by subtracting the test score from the retest score. In the initial test series only three *Ss* responded in a consistent direction to the six presentations of the gray stimuli. Thirteen *Ss* made no responses and the rest varied in making from one to five responses. In the retest series three *Ss* in Groups I and II made a single response in the opposite direction to that predicted and the rest of the *Ss* responded in the direction of mediational training if they responded at all.

Results

The mean frequency of response in the direction of white or black for each stimulus in the test and retest series is indicated in Table 2, for all groups. The difference scores were subjected to an analysis of variance with results as shown in Table 3. It may be noted that none of the interaction variances was significantly larger than the within-groups variance, thus permitting the use of the within-groups variance as the estimate of error (**8**). Bartlett's (**4**, pp. 195–197) test was used to test the hypothesis of homogeneity of variance and the resultant chi square was not significant.

TABLE 2

Mean Response to Gray in Test and Retest Series and the Difference Scores

(Four *Ss* in each group)

Groups	MA 1			MA 2		
	Test	Retest	Difference: Retest-Test	Test	Retest	Difference: Retest-Test
I_A	.25	5.50	5.25	−2.50	5.75	8.25
I_B	.25	5.75	5.50	− .75	5.50	6.25
II_A	− .50	2.50	3.00	−1.50	4.50	6.00
II_B	1.00	4.00	3.00	− .25	4.75	5.00
III_A	−1.00	3.00	4.00	−1.00	.50	1.50
III_B	− .75	−1.75	−1.00	− .75	2.25	3.00

TABLE 3

ANALYSIS OF VARIANCE OF DIFFERENCE SCORES

Source of Variation	*df*	Mean Square	*F*
Conditions (Groups A & B)	1	13.02	1.65
Treatments (Groups I, II, III)	2	78.90	9.98**
Levels (MA I & II)	1	35.02	4.43*
C × T	2	1.64	—
C × L	1	3.52	—
T × L	2	3.14	—
C × T × L	2	22.40	2.83
Within groups	36	7.90	
Total	47	116.59	

* Significant at the .05 level.
** Significant at the .01 level.

The difference in difference scores for the A and B groups was not significant indicating that the magnitude of change was approximately the same irrespective of the direction of change.

The mean differences between the test and retest scores for the treatments were 6.31 for Group I, 4.25 for Group II, and 1.875 for Group III, when subgroups were combined. As shown, the *F* for these treatment differences was highly significant. A test of the differences between the means for Groups I and II, and Groups II and III, produced *t*'s of 2.06 and 2.38 ($p < .05$), respectively. The within-groups variance was used as the error estimate, allowing 36 *df*. The hypothesis that the Group III scores did not differ significantly from chance was tested and a *t* of 4.00 ($p < .02$) was obtained.

The difference between the higher and lower MA groups was significant ($.05 > p > .01$). On the assumption that some *S*s in the higher MA group might formulate verbal statements with respect to the gray stimulus during mediational training, and assuming these verbal statements might occur in the retest situation to aid in mediating the appropriate response, it was predicted that there would be real interaction between the different treatments and mental age, the mentally older *S*s demonstrating more mediated generalization in Groups II and III. Contrary to this prediction, however, the treatment by levels interaction variance was not significantly larger than the within-groups variance.

DISCUSSION

The results of the present study have clearly demonstrated that response mediation can occur in young children and that motor responses as well as verbal responses are effective in the mediational role.

That the verbal group (Group I) would perform significantly better than the motor group (Group II) was expected on the assumption that auditory stimulation provided by verbal responses increases the number of internal stimuli as well as provides more distinctive cues. If one accepts this assumption, it would be presumed that the mediational stimuli provided by the verbal responses were nearly optimal in terms of distinctiveness. If this is the case, the relative performance of the motor group is especially striking, since the stimuli resulting from the simple turning of the hand 45° to the left or right would not be expected to produce particularly distinct stimulation and therefore other more discrete motor responses such as the movement of the left and right hands, or one hand and one foot, when used in a mediational role, might more closely approach the results obtained with the verbal responses.

If one further assumes that response mediation depends on a stimulus-response connection between the response-produced stimuli and the instrumental response, then one is obliged to propose some explanation for the results obtained from Group III (Control) since this group was not provided with the opportunity to build up this S–R connection. Although the performance of this group was relatively poor, the mean difference score was significantly larger than one would expect from chance. It had been predicted that spontaneous verbalizations might occur during the mediational training which could later serve in a mediational capacity. If this were the case, it was expected that since verbal behavior is related to mental age, spontaneous verbalizations would be more likely to

occur in the higher mental age groups. On the assumption that the performance of the higher MA *Ss* in Groups II and III would be increased relatively more than Group I which was given the appropriate mediational cues, one would expect relatively less difference in performance between the treatment groups in the higher MA level than between the treatment groups in the lower MA level. This effect then would show up in the treatment by levels interaction. However, since this interaction was not significant one may assume that either verbalization was occurring equally in both MA levels, or that the mediation was resulting from the continuation by *S* of a slight turning response which provided differential pressure on the fingers, or slight muscular tension which produced sufficient stimulation to serve in a mediational capacity.

The significance of the difference between the two MA levels appears to reflect a general superiority in performance on the part of the older *Ss*. It was noted in the experimental situation that the younger *Ss* often became concerned with such things as the number of marbles they had obtained or irrelevant aspects of the situation. Such behavior tended both to lower the number of responses and increase the number of errors, thus causing a decrease in the general performance level of the younger *Ss*.

In general the present investigation has also demonstrated that young children are capable of performing fairly complex tasks when instructions are given to point out the relation of the stimuli to the response and when sufficient motivation is provided. This fact, along with the knowledge that motor responses can serve in a mediational capacity, is particularly helpful. If, from further research of this sort, we can conclude that such responses parallel verbal responses in their general effects, then our interpretation of verbal behavior may be simplified. The motor response then offers an excellent opportunity to study the effects of mediational responses on behavior without incurring many of the complications involved in the use of verbal responses.

Summary

Forty-eight preschool children were trained to respond differentially to white and black stimuli. After meeting a criterion the generalization of the responses to gray was tested and found to be practically absent. Following this, Groups I and II acquired verbal and motor mediational responses, respectively, with subgroups which learned the same response for either black and gray or white and gray and these responses were then incorporated with the response previously learned. Group III learned the same motor mediational response as Group II but was not permitted to use this response during the retraining of the original response. Again the generalization to gray was measured and the data analyzed as difference scores between the first and second series. The verbal mediational response was found to produce significantly greater generalization than the motor response, but the motor response in turn produced relatively good generalization. That is, consistent responses were made to the gray stimuli in the retest session as the result of the inclusion of a mediational response with the original response learned to black and white.

Since Group III was a control group, the mediational response was not continued after the mediational training and consequently this group demonstrated very poor mediation. The difference scores, however, differed significantly from chance. This was explained by assuming that either spontaneous but subvocal verbalizations or covert turning responses were serving to mediate the generalization of the initial response.

The mediational process was assumed to depend on a stimulus-response connection between the stimuli produced by the mediational response and the instrumental response. Thus, when the mediational response which was trained to gray was the same as that trained to white, and when the mediational response to white was made prior to the pushing response to white for a number of trials, it was assumed that a stimulus-response connection was built up which later elicited the pushing response when this

mediational response was made to the gray stimulus.

(Received September 2, 1952)

References

1. Alberts, E., & Ehrenfreund, D. Transposition in children as a function of age. *J. exp. Psychol.*, 1951, **41**, 30–38.
2. Birge, J. S. The role of verbal response in transfer. Unpublished Ph.D. thesis, Yale Univer., 1941.
3. Dollard, J., & Miller, N. E. *Personality and psychotherapy.* New York: McGraw-Hill, 1950.
4. Edwards, A. L. *Experimental design in psychological research.* New York: Rinehart, 1950.
5. Gagné, R. M., & Baker, K. E. Stimulus predifferentiation as a factor in transfer of training. *J. exp. Psychol.*, 1950, **40**, 439–451.
6. Hull, C. L. *Principles of behavior.* New York: D. Appleton-Century, 1943.
7. Kuenne, M. R. Experimental investigation of the relation of language to transposition behavior in young children. *J. exp. Psychol.*, 1946, **36**, 471–490.
8. Lindquist, E. F. *Design and analysis of experiments in psychology and education.* (Mimeographed preliminary copy prepared at the State Univer. of Iowa, 1950.)
9. Lumsdaine, A. A. Conditioned eyelid responses as mediating generalized conditioned finger reactions. *Psychol. Bull.*, 1939, **36**, 650. (Abstract.)
10. Pyles, M. K. Verbalization as a factor in learning. *Child Developm.*, 1932, **3**, 108–113.
11. Rossman, I. L., & Goss, A. E. The acquired distinctiveness of cues: the role of discriminative verbal responses facilitating the acquisition of discriminative motor responses. *J. exp. Psychol.*, 1951, **42**, 173–182.
12. Shipley, W. C. Indirect conditioning. *J. gen. Psychol.*, 1935, **12**, 337–357.

Journal of Experimental Psychology
Vol. 45, No. 5, 1953

REINFORCEMENT AND "HYPOTHESIS" IN THE DISCRIMINATION BEHAVIOR OF CHIMPANZEES

H. W. NISSEN, BILLEY LEVINSON, AND J. W. NICHOLS

Yale University and Yerkes Laboratories of Primate Biology

One way of differentiating among theories of discrimination learning is on the basis of which one (or more) of several features of the stimulating situation are assumed to determine the response. Among these features are (*a*) the relation between the choice objects, (*b*) the configuration of two (or more) choice objects, and (*c*) the absolute excitatory strengths of the single objects.

The so-called relational theory holds that the ratio of excitation between the two objects—the "direction" of change (e.g., lighter, darker, smaller, larger) from one object to the other—provides the cue (**2**). It is especially clear in the case of successive discrimination that this theory assumes the relating of a present sensory event with the trace or memory of a previous one. The "configurational theory" (**1**) assumes that the undifferentiated pattern or Gestalt formed by the two (or more) objects is the effective cue. This theory necessarily describes the learned response in terms of differential muscle movements rather than as approach to a given part of the situation. The "absolute theory" (**5**) holds that the greater net excitatory strength of one object determines the response to it.

Solution of the discrimination problems described below on the basis of one-object patterns is possible. Each positive and each negative object used in training is a unique combination of sensory characteristics. According to the "absolute theory" one should expect nonarticulate organisms to respond to these simpler one-object patterns rather than to the more involved patterns formed by two stimulus plaques ("transverse patterning" **6**). However, the responses of our chimpanzee *S*s in subsequent tests, when they were confronted with new combinations of the objects used in training, show that choices were determined by both objects rather than by only one of them. These results are inexplicable on the basis of any one of the aforementioned determinants, but may be accounted for by assuming that *S*s responded to relational as well as absolute cues. The test responses reported below also demonstrate a marked tendency by chimpanzees to initiate their own "hypotheses" or responsiveness to certain stimuli, when no differential reinforcement has been given to those stimuli.

Design of the Experiments

An example of the training-testing design, used in all three experiments, is diagramed in Fig. 1. In a two-choice discrimination problem, a large white square (*W*) is paired with a small white square (*w*). Choice of *W* is always rewarded, choice of *w* is never rewarded. In other trials a large black square (*B*) is paired with a small black square (*b*), and in this case *b* is consistently rewarded, *B* is never rewarded. This training is continued until there is consistent choice of *W* in the first pair, and of *b* in the second pair. The performance may be verbalized in the formula: if white, choose large; if black, choose small. In subsequent double-baited test trials, the following novel combinations of stimulus objects are presented for choice: *W* vs. *B*; *b* vs. *w*; *W* vs. *b*; and *w* vs. *B* (see Fig. 1). What responses are to be expected in these tests?

(*a*) In the first two test pairs (which we shall call the positive-negative or P–N tests) a definite prediction follows from the absolute theory: *W* and *b* having been consistently rewarded, *B* and

w never having been rewarded, *S* should choose *W* and *b*.

(*b*) In the third and fourth test pairs (called the positive-positive or P–P, and negative-negative or N–N tests, respectively), there seems to be no basis for prediction. In terms of single-object patterns, large-plus-white and small-plus-black have both been rewarded; small-plus-white and large-plus-black have never been rewarded. In terms of component aspects (white, black, large, small), and of the relations between these (lighter, darker, smaller, larger), there likewise has been no differential rewarding which would favor one over the other choice object. The only bases of prediction would be a known innate preference for one cue or cue combination over others, or differential amounts of practice on the two training pairs.

In the work reported below, this experimental design was carried out with seven chimpanzees as *S*s. The number of basic (or trained) habits was two (Exp. I), four (Exp. II) or eight (Exp. III), and the corresponding number of test pairs (new permutations of the original stimulus objects) was 4, 24, and 112, respectively. The particular visual aspects used for the positive and negative objects varied among *S*s, as will be described. The significant findings are (*a*) that in the P–N tests there was only a 68% choice of the rewarded object, and (*b*) in the P–P and N–N tests some one of the nondifferentially rewarded stimuli elicited choice with reliably above-chance frequency.

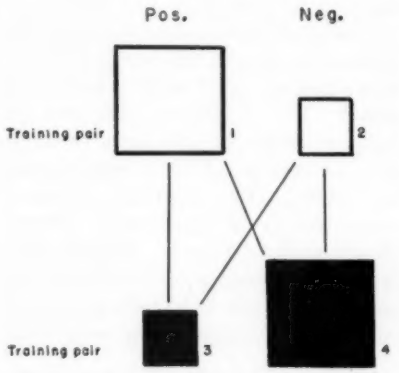

FIG. 1. Paradigm of the design of Exp. I, II, and III. Showing the two training pairs in horizontal rows (No. 1 vs. No. 2, and No. 3 vs. No. 4), and the test pairs: P–N test pairs (No. 1 and No. 4; No. 3 and No. 2) indicated by diagonals; P–P (No. 1 and No. 3) and N–N (No. 2 and No. 4) test pairs indicated by vertical lines.

General Procedure

Each *S* was trained in a "conditional discrimination" of the type exemplified in Fig. 1. A two-choice discrimination apparatus, previously described (**3**), was pushed up to the living cage of *S*. Reaching through the 2-in. mesh partition, *S* could displace the stimulus object and obtain a small piece of fruit in the food well under the correct object. Fifty trials were given in one session. The correction technique was used, *S* being forced to wait a few seconds after error before being permitted a second choice. Each habit, or pair of stimulus objects, was first trained to a criterion of ten successive errorless trials before the next habit was introduced. Subsequently there were more rapid shifts from one habit to another (cf. **3**), until the two, four, or eight habits comprising the conditional discrimination were being presented in random sequence. Training was continued until a criterion of not more than one error in two sessions (100 trials) was attained with random presentation of the several pairs of stimulus objects. Thereupon the testing sessions, which provide the main data of this report, were begun.

Each testing session began with ten randomized trials of the basic habits above described, and successive test trials were separated by four trials on the basic habits. In a total of 50 trials, therefore, only 8 were test trials. It was hoped in this way to make each test trial as "independent" of other tests as possible. The tests were of the P–N, P–P, and N–N kind as described in the legend of Fig. 1. They were given in the order: P–N, P–P, P–N, N–N. Chance left-right sequences, balanced in respect to both the tests and the intervening basic-habit trials, were used throughout. In all test trials both objects were baited, so there was never any "correction." This procedure was continued until the requisite number of tests had been accumulated. Further details, pertaining to the three main experiments, are given below.

Experiment I

Procedure.—The basic conditional discrimination consisted of two habits (as in Fig. 1), but the particular cues marking the positive and negative objects were different for each *S*. The cue combinations used are shown in Fig. 2. Size, color, and form were each used twice as intratrial variable (differentiating the two objects of a pair) and as intertrial variable (differentiating one pair of objects from the other). It should be noted that the intratrial or intrapair variable is always the same in both habits learned by a given *S*. The rewarded aspect of that variable (e.g., size), however, is different in the

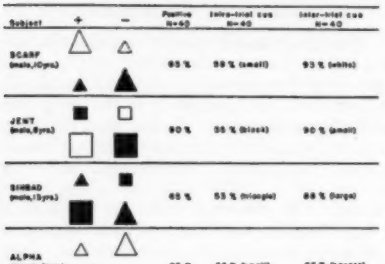

FIG. 2. Training objects used in the two basic discrimination habits for each of six *Ss* in Exp. I. The positive (rewarded) object is on the left. Percentages show test results: Responses to Positive in the P–N tests; responses to the preferred aspect of the intratrial variable (indicated in parentheses) in the P–P and N–N tests; and responses to the preferred aspect of the intertrial variable in P–P and N–N tests. (There were only 20 P–P and 20 N–N test trials, but all 40 trials presented both the inter- and intratrial variable.)

two habits (e.g., large in one case, small in the other). Exclusive of the 100 criterion trials, the learning scores for mastery of the two basic habits were as follows:[1] Scarf, 1400 trials; Jent, 1350; Sinbad, 1550; Alpha, 650; Verb, 200; and Dehn, 650.

Results.—The three columns on the right of Fig. 2 summarize results of the tests. There were 40 P–N test trials, 20 pairings of objects No. 1 and No. 4 (see Fig. 1) and 20 of objects No. 3 and No. 2. Although No. 1 and No. 3 had always been rewarded, No. 2 and No. 4 never, there was an average choice of No. 1 and No. 3 of only 69%. The difference between this score and that achieved on the 1920 interpolated trials with the two basic habits (95%) is highly significant ($p < .001$). The difference between the highest individual score (Jent, 80%) on the P–N tests and his score of 96% on the 320 interpolated training trials is significant at the .01 level. The group score (69%) is also significantly above chance ($p < .001$), and the lowest individual score (65%) is above chance at the .03 level of confidence. The results thus indicate that response to the rewarded object was affected significantly, but was not entirely determined, by the object with which it was paired.

The second column shows the percentage of choices of the preferred aspect of the intratrial variable in the 20 P–P and 20 N–N tests (pairing of objects No. 1 and No. 3, and of No. 2 and No. 4, respectively). Since both P–P and N–N test pairs presented this variable, there were 40 opportunities for choice. The preference for one or the other aspect (as shown in parentheses) exceeded the chance value of 50% by from 3 to 18%.

The third column shows comparable figures for the intertrial variable which also appeared in all 40 P–P and N–N tests.[2] The preference for one or another aspect of this variable deviates an average of 29% from chance ($p < .001$). The lowest individual score (65%) is significantly above chance at the .03 level. Whereas differential rewarding of the single objects used in the P–N tests produced only a 69% choice of the reinforced

[1] Left-right sequences which promote spatial alternation were inadvertently used in the early training of Scarf, Jent, and Sinbad; this accounts for their slow learning.

[2] The intertrial variable also appears in the P–N tests, but in these tests any preference in respect to the intertrial variable might be obscured by the presumably prepotent influence of differential rewarding. Actually, the percentages for the intertrial variable are almost the same when the P–N data are included as when they are omitted, as in Fig. 2.

object, a nondifferentially reinforced aspect of the intertrial variable effected a 79% preference for one of its alternative aspects.

Experiment II

Procedure.—After completing the tests of Exp. I, three of the same *S*s were trained in two additional habits, introducing a second intertrial variable (but in each case retaining the previously used intratrial variable). The total of four habits were mastered to a criterion of not over seven errors (not over three errors on any one habit) in 200 consecutive trials with a random sequence of the four pairs of objects. Exclusive of previous training (in Exp. I) and of the criterion trials, Alpha and Dehn each required 950 trials, and Verb required 2000 trials, for this basic training. As shown in Fig. 3, size, form, and color each appeared once as intratrial variable and equally often as intertrial variable.

Results.—The data of Fig. 3 show the same trends as those of Exp. I. The positive object in the P–N tests was again chosen in only 69% of the trials, significantly above chance ($p < .001$) and significantly different from the score of 98% (42 errors) on the 2880 interpolated trials with the four basic habits ($p < .001$). Preference for one or another aspect of the intra- and intertrial variables was less pronounced than in the previous experiment. The average of the maximum choice by each *S* for one of these nondifferentially rewarded aspects (62, 73, and 71%) was 69%, significantly above chance ($p < .001$). The lowest of these individual scores, 62%, is also significantly above chance ($p < .02$).

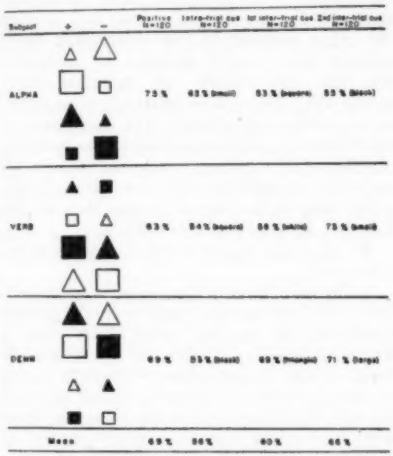

Fig. 3. Training objects used in the four basic discrimination habits for each of three *S*s in Exp. II. Test scores are in terms of percentage of responses (*a*) to the rewarded object, (*b*) to the preferred aspects of the intratrial and intertrial variables.

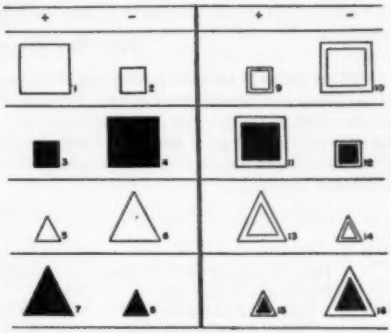

Fig. 4. Pairs of training objects used in the eight basic discrimination habits of Exp. III, one *S*, Frank, 12 yr.

Experiment III

Procedure.—The data of Exp. III were, chronologically, the first ones obtained in this study. Learning the eight basic discriminations has been reported previously (**3**); the present data are those mentioned in a footnote to Table 2 of that publication. The objects used in training are shown in Fig. 4, and the method of combining these for the P–N, P–P, and N–N test-pairs is described in the legend of Table 1.

The procedure of the testing sessions differed from that described above in that the number of "basic habit" trials interposed between successive test trials was only two, instead of four. Two series of tests were given, these being separated by an interval of almost one year. In each case, the eight basic discriminations were being performed errorlessly, when presented in random

TABLE 1

TEST RESULTS FOR EXPERIMENT III

(Test scores in terms of percentage of responses to the positive object and to the preferred aspect of the intra- and intertrial variables. Fifty-six P–N test-pairs were formed by combining every odd-numbered object with every even-numbered object (e.g., 1 and 4, 1 and 6) except the one constituting the training pair (e.g., 1 and 2). Twenty-eight P–P test-pairs were formed by all combinations of odd-numbered objects. Twenty-eight N–N test-pairs were formed by all combinations of even-numbered objects.)

	1st series	2nd series
Positive object in P–N tests	56%*	56%*
Dominant aspect of:		
Intratrial variable	94% small**	98% small**
First intertrial variable	53% black	51% black
Second intertrial variable	52% triangle	53% triangle
Third intertrial variable	59% margin absent	52% margin present

* Significantly different from the score of 90 per cent on the interpolated training trials, $p < .001$. The two series combined are significantly above chance at the .02 level.
** Significantly above chance, $p < .001$.

order, just before the testing began. In the first series, each P–N pair was presented four times, each of the other pairs twice. This gave a total of 336 test trials. In the second series, each test pair was presented twice, a total of 224 tests.

Results.—The test results are summarized in Table 1. This *S* responded to the previously rewarded member of the P–N pairs in only 56% of the tests. When the two objects of the P–P and N–N tests differed in respect to size, the intratrial variable, the smaller one was chosen in 94 to 98% of the trials. No pronounced preference for either aspect of the three intertrial variables was shown, except perhaps in respect to the absent margin in the first series. (An analysis of responses in terms of position shows 53% to the left in the first series, 54% to the right in the second series.)

DISCUSSION

Results of the P–N tests; what cues are reinforced?—The failure of *S*s of this experiment to respond consistently to the rewarded object versus the nonrewarded object shows that the "absolute theory" is by itself inadequate to explain the behavior of these animals. Although approach to the former, nonapproach to, or avoidance of the latter, provides an effective and apparently simple basis of solution, it is clear that the chimpanzees were responding to the particular combinations of choice objects used in training. Spence's statement that "substitution of some other form for the negative stimulus does not interfere in any manner with continued response to the positive stimulus" (6, p. 92) does not hold for the conditions of this experiment.[3] A change in the "irrelevant context" of the rewarded object, i.e., altering the other member of the object-pair, could account for a small decrease in accuracy, but does not explain the reduction from nearly 100% (on the interpolated training trials) to an average 68% choice. The significance of the P–N test data is emphasized by the fact that the degree of maximum preference shown by each *S* for a nondifferentially rewarded stimulus aspect was, on the average, greater than for the differentially rewarded alternatives.

We may now inquire whether other formula-

[3] Although Spence's exposition is quite condensed and does not specifically cover the present conditions, its meaning seems clear: when a single reinforced cue suffices, as in the triangle versus circle discrimination, there will be no patterning; where one-object or "local" patterning suffices, as in successive discrimination, there will be no two-object or "transverse" patterning; the latter occurs only when the cues deriving from each object singly are inadequate for solution. The problems of the present study can be solved by response to one-object patterns; nonarticulate organisms therefore should not have recourse to more complicated transverse patternings.

tions of discrimination learning are more successful in accounting for these results. It seems clear that, in spite of the theoretical adequacy of one-object patterns, both objects were somehow effective in determining the responses of our subjects. The "configurational theory" of Gulliksen and Wolfle (1) assumes that left- and right-going responses are learned to the undifferentiated two-object patterns. Since these patterns were no longer present when the objects were presented in new combinations, this theory would predict chance response in the P–N tests. But the data show an above-chance response to the rewarded object. Even in Exp. III, the deviation of 56% from chance is significant at the .02 level of confidence. Whereas the "absolute theory" founders on the fact of less than near-perfect response to the positive object, the "configurational theory" stumbles on the over-50% response.

Solution of the conditional discrimination problem may be based on responsiveness to a combination of relations. Approach to "larger plus lighter" and "smaller plus darker" solves the problem of Fig. 1. One of these relations (size) is provided perceptually, the other one (brightness) derives from a sensory event plus the memory or trace of a previous sensation. In this form, which we may call relational patterning, the "relational theory," like the "absolute theory," leads to the expectation of consistent response to the positive object. Both effective relations are, so to speak, available, although in the tests one component necessary to the size (rather than to the brightness) relation must be provided by memory.

Now responsiveness to a relation does not preclude concurrent responsiveness to the "absolute" properties of the components on which that relation is based. In approaching the large white square of Fig. 1, for instance, response may have been based partly on that square being larger than the other one, partly on its specific dimensions and brightness. In the subsequent tests (No. 1 vs. No. 4) the former cue was absent, the latter cues were still present. The combination of cues for approach to the large white square during training was (*a*) absolute size, (*b*) relative size, (*c*) brightness. In the tests, the second of these (*b*) dropped out and *S* had to depend on the combination of the remaining two, (*a*) and (*c*). It has been shown (4) that when the original two differentiating cues of a discrimination habit are reduced to one, accuracy of response drops below 100% but remains well above 50%. The present results would thus be explicable by assuming that in the original training these chimpanzees were depending on both the relational and absolute sizes of the stimuli, and that elimination of one of these cues resulted in a roughly proportionate decrease in accuracy. In Exp. III, the long experience (over 5000 trials) with the same intratrial variable (size) may have resulted in almost complete dependence on the relational cue.

Results of the P–P and N–N tests; indeterminism of "hypotheses."—In seven cases out of ten (Exp. I, II, and III) one of the nondifferentially rewarded stimulus aspects in the P–P and N–N tests elicited a higher frequency of choice than did the positive (versus negative) object in the P–N tests. (That is, the percentage in the "Positive" column is lower than the highest percentage in the other columns in all but three instances.) The average frequency of response to the most preferred aspect was 77% (as compared to 68% for the positive object in the P–N tests). We may inquire whether this finding is explicable in terms of innate or experiential factors.

As may be seen in Fig. 2 and 3 and Table 1, no one variable (size, color, form), and no one aspect of a variable (e.g., small, large), elicited consistent preference by all *S*s. The average frequency of choice for the dominant aspect of size (whether this was "small" or "large") was 75%, of color 63%, of form 60%. In six cases out of eight (4 from Exp. I, 3 from Exp. II, and 1 from Exp. III), small was preferred over large; white, black, square, and triangle were each preferred in four cases out of eight. The average frequency of choice for each of these six stimulus aspects, when it was dominant or preferred, was as follows: small, 74%; large, 80%; white, 71%; black, 54%; square, 61%; triangle, 60%. Species-constant, innate preferences for (or dominance of) certain variables or stimulus aspects, therefore, do not explain the results of the P–P and N–N tests.

In Exp. I, the intertrial variable elicited a higher degree of preference than did the intratrial variable in six cases out of six; the average frequency for the dominant aspect of the former was 79%, for the latter 61%. In Exp. II the intratrial variable gave the higher frequency of choice in 1 out of 3 cases (Alpha). In Exp. III, the intratrial variable (size) showed a far greater differential choice than did any of the three intertrial variables. The P–P and N–N test results, therefore, cannot be explained consistently in terms of differences between the intratrial variable (where simultaneous comparison is possible) and the intertrial variable (the alternative aspects of which do not appear simultaneously).

Looking again at Fig. 1, we see that if more practice had been given on the first training pair (No. 1 vs. No. 2) than on the second (No. 3 vs. No. 4), "large" would have been reinforced more often than "small," and "white" more often than

"black." Differences in amount of training on each of the several habits learned by one *S could* explain differences in preference. Analysis of the protocols, however, reveals that for each *S* the number of training trials on the several habits was very nearly the same. The maximum difference, in Exp. I and II, between the least and most practiced training pair was 4.5%. More important is the fact that preferences shown on the tests were more often contrary to expectation based on such training differences than in accordance with them. Further analyses were made in terms of reinforcements minus nonreinforcements, and over-all accuracy scores. These also showed more instances contrary to, than consonant with, expectation. In Exp. III, first series, the four habits in which "small" was correct had 6% more training trials than the four in which "large" was correct. It will be recalled that "small" was overwhelmingly dominant over "large" in the tests (94%). But a 61% advantage in number of training trials on "margin absent" is associated with a preference of only 59% in the tests (this dominance being reversed in favor of "margin present" in the second series of tests). "White" had a practice advantage of 10% over "black," and yet black was dominant over white (53%) in the tests. "Square" was positive in 1% more training trials than "triangle," but the latter was slightly dominant in the tests. The preferences for nondifferentially rewarded stimulus aspects are therefore not accounted for by differences in amount of practice during training.

The possibility that unknown experiences provided a basis for the preferences shown in the P–P and N–N tests cannot, of course, be completely excluded. For the present, however, we are left with the unsatisfactory conclusion that the results were a function either of individually variable innate determinants, or of arbitrary, unpredictable whims or "hypotheses" of the individual *S*s.[4] It is possible that variations in the "emphasis" or "attention" given by *S* to one or another stimulus are responsible for the uneven distribution of choices. The complete determination of such variations may be beyond experimental control.

[4] To say that, since both objects were baited in the tests, the scores are a function of learning after "accidental" choices on the first few tests, does not essentially alter this conclusion. The arbitrariness is then merely confined to the early tests. Furthermore, the conclusion seems improbable: choice on the first test on which a given variable appeared corresponded to the over-all preference in all tests (percentages in Fig. 2 and 3 and Table 1) in only 60% of the cases.

Summary

Seven chimpanzees were first trained in conditional discriminations consisting of two, four, or eight habits. Each stimulus object was consistently rewarded or not rewarded. Subsequently the same objects were presented in novel pairings or combinations. The main results were as follows:

1. When a previously rewarded object was paired with a previously nonrewarded object, the former was chosen in only 68% (average) of the tests. It is suggested that the animals had been responding to both absolute and relational aspects of the stimuli, and that in the absence of the latter cue accuracy was correspondingly reduced.

2. When two previously rewarded objects, or two previously nonrewarded objects, were paired, some nondifferentially reinforced stimulus aspect, which varied among *S*s, was chosen in from 62 to 98% of the tests. No satisfactory explanation of these preferences was found.

(Received September 2, 1952)

References

1. Gulliksen, H., & Wolfle, D. L. A theory of learning and transfer. I. *Psychometrika*, 1938, **3**, 127–149. II. *Psychometrika*, 1938, **3**, 225–251.
2. Lashley, K. S., & Wade, M. The Pavlovian theory of generalization. *Psychol. Rev.*, 1946, **53**, 72–87.
3. Nissen, H. W. Analysis of a complex conditional reaction in chimpanzee. *J. comp. physiol. Psychol.*, 1951, **44**, 9–16.
4. Nissen, H. W., & Jenkins, O. W. Reduction and rivalry of cues in the discrimination behavior of chimpanzees. *J. comp. Psychol.*, 1943, **35**, 85–95.
5. Spence, K. W. The nature of discrimination learning in animals. *Psychol. Rev.*, 1936, **43**, 427–449.
6. Spence, K. W. The nature of the response in discrimination learning. *Psychol. Rev.*, 1952, **59**, 89–93.

Journal of Experimental Psychology
Vol. 45, No. 5, 1953

APPROACH-AVOIDANCE IN AN AMBIVALENT OBJECT DISCRIMINATION PROBLEM[1]

ROBERT THOMPSON[2]

Yerkes Laboratories of Primate Biology, Orange Park, Florida

In a previous study, Nissen (4) obtained data from the conventional two-choice situation which clearly demonstrated the inadequacy of describing the learned discriminative response in terms of specific muscular movements. Chimpanzees trained on a visual discrimination problem with the stimulus plaques in one spatial arrangement (horizontal plane) exhibited almost perfect transfer when the stimuli were presented in an altogether different spatial arrangement (vertical plane). If the *Ss* had learned only "left-going" and "right-going" responses, differentially elicited by discrete configurations of stimuli, as proposed by Gulliksen and Wolfle (1), this change in the spatial relationship of the discriminanda would have completely disrupted the habit. The fact that the habit was maintained after shifting from the horizontal to vertical plane—or vice versa—indicates that the *Ss* had learned to approach the positive object and/or avoid the negative object.

Although an "approach-avoidance theory" as contrasted with a "learned-movement theory" receives considerable support from the simple two-choice problem, it is conceivable that discriminative situations of greater complexity may demand modes of response incompatible with those described in approach-avoidance terms.[3] What we have called an "ambivalent object" discrimination problem may represent one of these complex situations. In the example suggested by Spence (6), three different figures are paired in all possible combinations so that each figure is both positive (rewarded) and negative (unrewarded): triangle (+) versus circle (−); square (+) versus triangle (−); circle (+) versus square (−). The task of *S* is always to choose the positive figure when all three combinations are presented in random order. This problem differs from others in that no stimulus object or one-object pattern is consistently reinforced or nonreinforced, and consequently its solution requires a response which is determined by the total stimulus situation. Solution of the problem as proposed by Spence may be explained in terms of either (*a*) movement learning or (*b*) approach-avoidance learning. According to the movement-learning conception the three object-pairings shown above, together with their mirror images, provide a total of six unique two-object configurations, to which *S* might learn left- and right-going responses. Thus triangle-circle elicits a left-going response, whereas circle-triangle elicits a right-going response.

[1] The writer wishes to acknowledge the generous assistance of Dr. H. W. Nissen in the planning of this study and in the preparation of this manuscript.

[2] Now at the University of Texas.

[3] Weise and Bitterman (7) believe that the "successive problem," which they found to be learned more readily than the "simultaneous problem," represents a situation which cannot be appropriately described in terms of approach-avoidance. However, Nissen (5) has shown that the results obtained by these investigators are explicable in approach-avoidance terms, and consequently, that the assumption of learned movements is not necessary. Furthermore, Spence (6) reports several studies which contradict the findings of Weise and Bitterman.

Alternatively, *S* may learn to approach the triangle if it is paired with a circle, to avoid the triangle when it is paired with a square. Described thus, in approach-avoidance (or approach-non-approach) terms, the problem involves what Spence (6) calls "transverse patterning" and what Nissen terms "ambivalent cues" (3) or "conditional discrimination" (4).

A crucial test as to which of these response descriptive categories is applicable to the ambivalent object discrimination problem may be had by following the design employed by Nissen. If *S*s are trained on all three habits with the objects in one plane (i.e., horizontal), and are then tested with the objects in a different plane (i.e., vertical), the movement-learning theory would predict no transfer whatsoever, whereas the approach-avoidance theory would predict a high degree of transfer.

It was the purpose of this study to determine (*a*) whether chimpanzees are capable of learning the problem, as Spence has predicted, and (*b*) whether the movement or approach-avoidance formulations more adequately account for the performance.

Method

Subjects.—Five preadolescent male chimpanzees, their ages ranging from 5 to 7 yr., were used as *S*s. Subjects Dehn and Verb had previously participated in an experiment of this nature involving a shift of the relevant cues from one positional relationship to another (4), while Frans was formerly used as an *S* in a "learning set" program (2). Alan and Ked had been subjected to ablation of prefrontal cortex at the age of two months, and Ked had undergone also a partial ablation of parietal cortex; these operated animals had had limited discrimination experience prior to the beginning of this study. All *S*s were housed in individual home cages throughout the experiment, and work sessions were given in the morning hours before regular feeding.

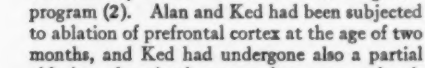

Fig. 1. Stimulus patterns employed in establishing Habits I, II, and III, respectively.

Apparatus.—The apparatus was essentially the same as that used by Nissen (4). Briefly, the stimulus holder consisted of two compartments into which stimulus plaques could be inserted. The holder was so constructed that it could be employed either in the horizontal plane (plaques presented in a right-left relationship) or in the vertical plane (plaques presented in an up-down relationship). The displacement of a plaque (pushing it backwards into the holder) when it appeared in either plane revealed a food well.

The stimulus holder was mounted on a horizontally sliding tray which was enclosed in a wooden superstructure containing a vertically sliding opaque screen. During a trial, the screen was raised and the tray moved forward to a position where *S* reached through the wire mesh of its cage to displace one of the plaques. If the choice was correct, *S* obtained the reward (a small piece of fruit) from the food well under the displaced plaque. The tray was then drawn back, the screen lowered, and preparations were made for the next trial.

Three plaques, 5-in. square, were used as test stimuli: plain red, diagonal black and white stripes (1-in. wide striations), and a plain blue upon which was centered a disc of ¼-in. plywood, painted yellow, and 2 in. in diameter. It was believed that these stimulus plaques—each plaque different from the others in several visual aspects—might facilitate learning.

Habituation trials.—Training was preceded by 30 habituation trials in which all *S*s were required to displace a single "neutral" tan plaque to obtain a reward. The stimulus holder was in the horizontal plane for 15 trials, and in the vertical plane for 15 trials.

Training trials.—For all *S*s the three habits to be mastered concurrently are shown in Fig. 1. Thirty trials were given each day and the non-correction method was used throughout. Frans, Dehn, and Ked were given training in the vertical plane, while Alan and Verb received training in the horizontal plane. The training procedure was not rigidly fixed for all *S*s, but was varied according to the speed with which each *S* learned the various habits. In general, Habit I (red vs. striations) was learned to a criterion of ten consecutive correct responses, and subsequently Habit II (disc vs. red) was learned to the same criterion. Thereupon, Habits I and II were each relearned to a gradually decreasing criterion until a high degree of accuracy was attained when the two habits were performed in random order. Habit III was then established while retention trials were provided for Habits I and II. At this point, each pair of objects was given in blocks of three to five trials; the order

TABLE 1

LEARNING AND TRANSFER SCORES FOR EACH SUBJECT

S	Learning (Precriterion)		Criterion (60 Trials)	Transfer (60 Trials)	Error Difference
	Trials	Errors	Errors	Errors	
Alan (H to V)*	358	69	4	6	2
Verb (H to V)	707	180	5	8	3
Dehn (V to H)**	766	228	6	4	−2
Frans (V to H)	1031	327	6	6	0
Mean	715.5	201.0	5.2	6.0	0.8

* Trained on horizontal, tested on vertical.
** Trained on vertical, tested on horizontal.

in which they appeared was mixed. Upon reaching a high level of performance, the three pairs were presented in a prearranged "random" order, one particular combination not appearing over twice in succession. Training was continued until *S* made not less than 90% correct responses in two consecutive sessions (not more than six errors in 60 trials, with not more than three errors on any one habit).

After reaching the criterion, each *S* was given ten trials with the single neutral tan plaque in a spatial position to be used in the transfer test, the purpose of this being to reduce the "shock" of seeing the plaque holder in a different orientation.

Transfer test.—Two sessions of 30 trials each, in which the three combinations of stimulus plaques were presented in random order, but in a spatial relationship other than that used in original training, constituted the transfer test. The transfer test was instituted on the day following the last of the two criterion sessions. Except for the change in spatial relationship, the transfer procedure was the same as that used during the criterion of the original learning.

RESULTS AND DISCUSSION

Original learning of the ambivalent object discrimination problem.—Four of the five *S*s succeeded in mastering the ambivalent object discrimination problem. Ked, the animal with both prefrontal and parietal cortical ablations, failed to learn the problem within 2100 trials. Although Ked was able to master any two of the three habits, training on the third habit invariably disrupted performance on the two acquired habits. Various adjustments in procedure, such as changing to the correction technique, and shifting training from the vertical to the horizontal plane, did not result in any appreciable improvement. Ked is a very emotionally responsive animal and is inclined to cry upon making an error, a reaction which persisted even after the correction method was instituted. This factor as well as the brain damage may have contributed to his failure.

Table 1 summarizes the data obtained on the original learning and on the transfer test for the four remaining *S*s. Alan (prefrontal cortical extirpation) required the fewest trials and committed the fewest errors in reaching the criterion. Frans, on the other hand, exhibited the least efficiency in solving the problem, requiring over 1,000 trials to attain the criterion. It is interesting to note that of the four *S*s who learned the problem, Alan had the least discrimination experience prior to the beginning of this study, whereas Frans had the most. This suggests that previous experience with simple discrimination problems, especially in the "learning set" procedure, may have a detrimental or retarding effect upon the later learning of an ambiguous object discrimination problem which demands many repeated trials on a given set of discrimanda.

Performance on the transfer test.—The results of the transfer test clearly demonstrate that changing the spatial

relationship of the discrimanda failed to disturb efficiency of performance on the ambivalent object discrimination problem. For each *S*, the number of correct responses on the 60 transfer trials was very significantly above chance expectation ($p = .001$). The maximum number of errors made in the 20 transfer trials on any one habit was 4 (by Verb on Habit II, and by Frans on Habit III); 16 correct responses in 20 trials is significantly better than chance at the .01 level of confidence. That is, each *S* performed significantly above chance on each of the three habits constituting the 60 transfer trials. Furthermore, in no individual case was the total number of transfer errors significantly greater than the number of criterion errors.[4] Verb, who made the most errors on the transfer test (8 errors), was still performing at an efficiency that was not reliably inferior to that which he exhibited during the 60 criterion trials ($p = .50$).

The implications of the data are unambiguous: since a change in the spatial arrangement of the stimulus plaques did not disrupt subsequent performance, the learning of the ambivalent object discrimination problem could not have been based on the acquisition of appropriate muscular movements, each elicited by a different two-object configuration. The high degree of transfer exhibited by each *S* can only be explained if we assume that the *S*s had learned approach to a given stimulus plaque depending, or conditional on the plaque with which it was paired. It may be concluded, furthermore, that inapplicability of the approach-avoidance formulation to simple or complex discrimination problems has yet to be demonstrated.

[4] It should be mentioned that the errors committed within the 60 transfer trials were quite evenly distributed throughout the two sessions devoted to the transfer test. The average number of errors for all four *S*s was 3.2 in the first session, 2.8 in the second session of 30 trials. This suggests that the transfer trials represent tests of previous learning rather than new learning.

Summary

The present experiment showed that chimpanzees are capable of learning an ambivalent object discrimination problem. Three different stimulus plaques were paired in all possible combinations so that each plaque was positive when it was paired with one stimulus object, negative when it was paired with the other stimulus object. Four of the five *S*s succeeded in consistently choosing the positive plaque when the three combinations were presented in random order.

A change in the spatial arrangement of the discriminanda from the horizontal to the vertical plane—or vice versa—failed to disturb subsequent performance. This clearly demonstrates that the learned discriminative response is more appropriately described as approach-avoidance rather than in terms of learned movements.

(Received September 8, 1952)

References

1. Gulliksen, H., & Wolfle, D. L. A theory of learning and transfer. *Psychometrika*, 1938, **3**, 127–149.
2. Hayes, K. J., Thompson, R., & Hayes, C. Discrimination learning sets in chimpanzees. *J. comp. physiol. Psychol.*, in press.
3. Nissen, H. W. Ambivalent cues in discrimination behavior of chimpanzees. *J. Psychol.*, 1942, **14**, 3–33.
4. Nissen, H. W. Description of the learned response in discrimination behavior. *Psychol. Rev.*, 1950, **57**, 121–131.
5. Nissen, H. W. Further comment on approach-avoidance as categories of response. *Psychol. Rev.*, 1952, **59**, 161–167.
6. Spence, K. W. The nature of the response in discrimination learning. *Psychol. Rev.*, 1952, **59**, 89–93.
7. Weise, P., & Bitterman, M. E. Response-selection in discrimination learning. *Psychol. Rev.*, 1951, **58**, 185–195.

Journal of Experimental Psychology
Vol. 45, No. 5, 1953

THE EFFECTS OF SINGLE AND COMPOUND CLASSES OF ANAGRAMS ON SET SOLUTIONS[1,2]

IRVING MALTZMAN AND LLOYD MORRISETT, JR.

University of California at Los Angeles

It has been assumed by the authors (2) that the term *mental set* refers to behavior which can be accounted for in terms of the acquisition of habit strength by a class of stimuli for the elicitation of a particular class of responses. If this is the case, implications of the principle of habit strength and related principles of Hull's theory of behavior should hold for phenomena described as mental set. A previous study (2) tested an implication of the assumption that habit strength is a negatively accelerated growth function. An inverse relationship obtained between amount of training with incompatible anagrams and subsequent success in the solution of one of these classes of anagrams was in accord with these assumptions.

The present study is concerned with the corollary of Hull's theory which states that the habit strength acquired by different elements of a stimulus compound summate in the same manner as the increments in habit strength acquired by a given stimulus with consecutive reinforcements (1, ch. 13).

Relevant empirical evidence is Hull's (1) report of the results of a conditioning study. A light conditioned to the GSR elicited a mean response of 3.5 ml. A vibratory stimulus next conditioned to the GSR elicited a reaction of 3.6 ml. When both conditioned stimuli were presented simultaneously, the jointly evoked mean reaction was 4.4 ml. Similar results were obtained under a variety of different experimental conditions. According to Hull the shrinkage in response strength of the reaction elicited by the compound stimulus as compared with the component stimuli is given by the index $R_{1+2}/R_{1+}R_2$. In the present illustration its value was 62%. Under a variety of different experimental conditions the index varied from 60% to 66% indicating a shrinkage in response-evocation power of a compound stimulus of approximately ⅓. Although the jointly aroused response is greater than the response elicited by either stimulus alone, it is less than the arithmetic sum of the reactions elicited by the component stimuli.

The major corollary involved in the derivation of these findings was the one previously stated that the habit strengths acquired by different components of a stimulus compound summate in the same manner as the habit strengths acquired by a given stimulus with consecutive reinforcements. The growth of habit strength is exponential in both cases. An additional principle which may be of some importance in such situations is that of afferent neural interaction and the related effects of stimulus generalization. Finally, it must be assumed that habit strength and response strength are monotonically related.[3]

The purpose of this experiment was to investigate the following implications for mental-set phenomena of Hull's corollary concerning the summation of habit structures and related principles:

1. A compound class of anagrams unchanged during the training and test series would be followed by a greater percentage of set solutions than constant single classes of anagrams. This would be the case since the former condition would acquire the summated strength of two habit

[1] This study was aided by a grant from the Committee on Research, Graduate Council, University of California.

[2] Portions of this paper were presented at the 1952 Convention of the Western Psychological Association.

[3] A more rigorous and detailed derivation of these findings with related corollaries may be found in Hull (1, ch. 13).

structures acquired during the training and test series.

2. A compound class of anagrams constant during the training and test series would be followed by a greater percentage of set solutions than single classes of anagrams formed in the test series following training with compound classes. The latter condition would be followed by fewer set solutions since they would be a function of the strength of a single habit structure as compared with the summated habit structures in the constant compound condition.

3. A constant compound class of anagrams would be followed by a greater percentage of set solutions than a compound class formed in the test series following training with a single class. This follows from the assumption that the constant compound class would acquire summated habit structures during the training series while a single class of anagrams would not.

4. Addition of a second class of anagrams to form a compound class during the test series would be followed by an increased percentage of set solutions as compared with a constant single component class. This follows from the assumption that formation of the compound class in the test series would permit the summation of some habit strength acquired by the added component class of anagrams.

Method

Subjects.—Seventy volunteers from introductory psychology classes served as *S*s in the experiment. There were ten *S*s in each of the seven experimental conditions.

Procedure.—The training series for all conditions consisted of 20 five-letter anagrams having only one word as a solution which were typed in the lower case on white 3 × 5-in. index cards. A test series followed of 15 five-letter anagrams possessing several different words as solutions. Only one of the alternative solutions was a set solution related to the given class of training anagrams.

TABLE 1

Words Employed as Anagrams in the Practice, Training, and Test Series

Practice Anagrams: brick, chair, erase.

Nature training words		Order training words	
ponds	plant	queen	jelly
river	grass	camel	hands
brook	ferns	child	cigar
creek	stems	ships	cards
swamp	twigs	clock	books
water	seeds	tents	balls
rains	daisy	fresh	linen
snows	tulip	match	desks
storm	bulbs	towel	mouth
foggy	pansy	dozen	lends
Nature test words		**Order test words**	
stalk, talks		cause, sauce	
petal, plate, pleat, leapt		least, steal, stale, tales, slate	
weeds, sewed		march, charm	
marsh, harms		satin, stain, saint	
field, filed		baker, break, brake	
lakes, leaks, slake		board, broad	
coast, coats		broth, throb	
tides, diets, deist, edits		teach, cheat	
reefs, frees		trace, crate	
peaks, speak, spake		gates, stage	
dales, leads, deals		sword, words	
night, thing		lamed, medal	
sleet, steel		waits, waist	
range, anger		wrote, tower	
trout, tutor		worth, throw	

The different experimental conditions were obtained by varying the class of anagrams presented during the training and test series. In the nature-nature or N–N condition the solution of each of the anagrams in the training series was a word referring to some aspect of nature, for example, "river." The test anagrams had one solution referring to some aspect of nature and at least one other solution not directly or obviously related to nature. For example, the words "marsh" or "harms" would be solutions for the same anagram, but only the former would be a set solution. The different anagrams in this condition were in no consistent letter order or arrangement.

A second condition was the order-order or O–O condition. The training anagrams had only one possible solution and all were in the same letter order, 32145. The third letter in the word was the first in the anagram, and so on.

The anagrams were not consistently related to each other in any other obvious manner. Anagrams in the test series all had multiple solutions. The set solution was in the 32145 letter order.

A third condition was the order-nature—order-nature or ON–ON condition. Anagrams in the training and test series were the same as those in the N–N condition. In addition they were in the letter order of the O–O condition for the training series. In the test series the set solution was a nature word in the 32145 letter order. The alternative solutions in the test series were unrelated words in no consistent letter order. Thus this condition was the compound of the previous two conditions.

The remaining four conditions of the experiment were different combinations and permutations of the three conditions previously described. These conditions were order-nature–order (ON–O), order-nature–nature (ON–N), order–order-nature (O–ON), and nature–order-nature (N–ON). In the first three experimental conditions the set or sets established in the training series were utilized in the test series while in the last four experimental conditions a change in set or sets occurred between the training and test series. Words used to form anagrams for the different conditions are shown in Table 1. Most of the words were obtained from lists used by Rees and Israel (4).

A 2-min. time limit was enforced for each anagram in the training and test series of all conditions. During the training series, if *S* failed to obtain a correct solution in the allotted time, or gave an incorrect solution, *E* gave the correct solution. This correction technique was not employed during the test series. If *S* did not solve a test anagram within the allotted time or gave an incorrect solution, *E* presented another anagram without comment. Three practice anagrams were presented prior to the start of the experiment. They were in random letter orders and were not directly related to the class of nature anagrams. Instructions to *S*s were the same as those previously employed (2).

Results and Discussion

The mean percentages of set solutions in the test series obtained under the different experimental conditions are presented in Table 2. Bartlett's test gave a chi square of 13.2; $df = 6$, $p = <.05$. Since the variances were heterogeneous, an analysis of variance of the ranked data was conducted (5). A χ^2r of 18.0 was obtained; $df = 6$, $p = .007$. The ranked data of each condition were therefore compared with every other. The probabilities of the χ^2r comparisons for a two-tailed test are also shown in Table 2. The tabled probabilities should be halved for comparisons where the direction of the difference was explicitly predicted.

The results show that the constant compound condition ON–ON produced a greater percentage of set solutions than any other condition. The difference between the ON–ON and N–N conditions was highly significant. However, the difference between the ON–ON and O–O conditions only approached significance; $p = .10$ with

TABLE 2

Mean Percentage of Set Solutions and p Values for Independent χ^2r Comparisons*

	ON–N	O–O	O–ON	N–ON	N–N	ON–O	ON–ON
ON–N		.11	.20	.35	.20	.60	.005
O–O			.80	.60	.06	.80	.20
O–ON				.60	.06	.06	.60
N–ON					.06	.35	.03
N–N						.80	.002
ON–O							.60
Means	63.3	73.3	76.0	70.2	55.3	61.3	87.3
Variances	121.1	621.9	396.6	258.9	227.4	511.7	83.5

* All differences significant at the .06 level of confidence except between conditions N–ON and N–N were significant at or beyond the .05 level using the t test following the conversion of the data into angles by means of the arc sine transformation.

a one-tailed test.[4] It was also found that the constant component conditions differed in the percentage of set solutions they produced. The O–O condition produced significantly more such solutions than the N–N condition.

In accord with theoretical expectations the constant compound class of anagrams ON–ON was followed by a greater percentage of set solutions than the compound classes produced by addition of a class of solutions in the test series, O–ON, and N–ON. However, only the difference between ON–ON and N–ON was significant.

Confirmation of another implication of learning theory was found in the results of removal of one class of anagram solutions from the compound condition. This "subtraction" of a set decreased the percentage of set solutions as compared with the constant compound condition. The ON–ON condition was followed by a significantly greater percentage of set solutions than the ON–N and ON–O conditions.

The results also show that the addition of a second class of anagram solutions during the test series tends to increase the percentage of set solutions as compared with a single class of anagrams. However, only the difference between N–N and N–ON was significant; $p = .03$ with a one-tailed test. The difference between O–O and O–ON was not significant.

The results of the individual comparisons among the different conditions in general support the S–R interpretation of mental set presented here. The significant difference obtained between the O–O and N–N conditions, however, altered the conditions necessary for an adequate test of some of the implications of Hull's behavior theory for mental-set phenomena. The greater percentage of set solutions in the O–O condition as compared with the N–N condition may be due to differences in effective habit strength acquired by the two groups. Following successful solutions in both conditions there is an increase in reaction potential for the elicitation of the given class of responses by the given class of stimuli. In the case of order anagrams when a differential amount of habit strength has been established, a mental set has been acquired, for the arousal of a class of responses, any particular anagram belonging to this class can be solved without further selective learning. This does not occur with the class of nature anagrams. Although the habit strength for solutions to this class of anagrams is dominant, the solution of any particular anagram may still involve selective learning. In other words, a principle is involved in the solution of order anagrams. There is none in the solution of nature anagrams.

Another factor which may have been responsible in part for the smaller amount of effective or generalized habit strength acquired by the nature class of anagrams is the tendency for these words to fall into two distinct subgroups, words referring to bodies or forms of water and words referring to plant life. Competing responses

[4] If the percentages of set solutions in the ON–ON, O–O and N–N conditions are substituted in Hull's shrinkage index, a value of 67.9% is obtained. Rees and Israel (4) reported the results of an experiment on congruent sets, a condition similar to the compound ON–ON condition of the present study, and the results of experiments similar to the O–O and N–N conditions. Substituting the results of their three experiments in the shrinkage index 65.9% set solutions is obtained.

Although the value of the shrinkage index appears to be fairly constant, the possibility of its being an artifact of the use of percentages should not be overlooked. A great deal of experimental and theoretical work is necessary before the generality of the index can be adequately determined.

may have been aroused under these conditions. Evidence supporting the latter assumption was obtained in an experiment by the authors (3). When the nature class of anagrams was made more homogeneous, a significant difference between ON–ON and O–O conditions was obtained. Rees and Israel (4), however, did not obtain any difference in frequency of set solutions between nature and the particular letter order anagrams employed in the present study.

It was previously stated that the differences between ON–ON and O–O and between ON–ON and O–ON were not significant. An explanation for the failure to obtain a larger difference in these two cases is related to the significant difference obtained between the O–O and N–N conditions. The greater percentage of set solutions in the O–O condition indicates that a relatively great amount of effective habit strength for the arousal of set solutions was acquired here. Since the growth of habit strength is negatively accelerated, further increments in a compound contributed by solutions of the class of nature anagrams would be relatively small. The greater habit strength for order solutions tends to "overshadow" that for nature anagrams. In contrast the relatively low percentage of set solutions obtained in the N–N condition indicates that a relatively small amount of effective habit strength for the arousal of set solutions was acquired in this condition. Further increments in a compound contributed by solutions of the class of order anagrams would therefore be relatively large. The significant difference between N–N and N–ON and the absence of such a difference between O–O and O–ON is in accord with this interpretation. The significant difference between the ON–ON and N–ON conditions and the absence of a significant difference between ON–ON and O–ON conditions is likewise in accord with the above interpretation.

Interpretations other than the one presented here may be suggested to account for the data. One such alternative is that the significant differences obtained simply reflect the greater difficulty of anagrams in no consistent letter order as compared with the 32145 letter order. The significant difference between the O–O and N–N conditions as well as other differences obtained indicate that letter order is an important variable. However, a number of significant differences among the conditions are contrary to the assumption that the variable of letter order can account for all of the results obtained. For example, the ON–ON condition produced significantly more set solutions than the ON–O condition even though the letter order was the same in training and test series for both conditions. On the other hand, the letter order in conditions ON–O and N–N is different in both the training and test series, yet there is no significant difference in the percentage of set solutions obtained in the two groups. These results indicate that differences in letter order among the various conditions cannot completely account for the experimental results.

Another alternative interpretation is that an ON condition in the test series was generally followed by a greater percentage of set solutions than an O or an N condition because the compound condition provided two different ways of solving the same anagram with a set solution as compared with only one method in the single conditions. There was a greater probability of an *S*'s hypotheses giving set solutions in the compound condition than in the single conditions. The increased percentage of set solutions in the constant compound condition was not due to the summation of habit structures acquired during the training period. If this theory were correct, all ON conditions in the test series should have the same percentage of set solutions and all of these conditions should have a greater percentage of set solutions than single classes of anagrams in the test series. But this was not the case. For example, the constant compound condition ON–ON produced significantly more set solutions than the newly formed compound N–ON. This indicates that prior training with a single nature class of anagrams had a different effect from prior training with a compound class of anagrams. According to the present formulation the difference was due to the summation of habit structures during the training series that occurred in the constant compound condition.

Summary

Implications of Hull's principles of behavior concerned with the reaction evocation power of compound stimuli were investigated with respect to the tendency of single and compound classes of anagrams to arouse set solutions. The following were some of the results obtained:

1. A constant compound class of anagrams, ON–ON, produced a greater percentage of set solutions than a constant single class of order anagrams, O–O, and a single class of nature anagrams, N–N. However, only the latter difference was significant.

2. The constant single class of order anagrams produced significantly more set solutions than the constant single class of nature anagrams.

3. The constant compound class of anagrams produced significantly more set solutions than single classes of anagrams formed in the test series following training with a compound class.

4. The constant compound class of anagrams produced a significantly greater percentage of set solutions than a compound class formed in the test series following training with nature anagrams but not with order anagrams.

5. Addition of a constant letter order to the nature class of anagrams significantly increased the percentage of set solutions, but the addition of nature to letter order did not.

(Received August 14, 1952)

References

1. Hull, C. L. *Principles of behavior.* New York: D. Appleton-Century, 1943.
2. Maltzman, I., & Morrisett, L., Jr. Different strengths of set in the solution of anagrams. *J. exp. Psychol.*, 1952, **44**, 242–246.
3. Maltzman, I., & Morrisett, L., Jr. The effects of task instructions on solution of different classes of anagrams. *J. exp. Psychol.*, 1953, **45**, 351–354.
4. Rees, H., & Israel, H. An investigation of the establishment and operation of mental sets. *Psychol. Monogr.*, 1935, **46**, No. 6 (Whole No. 210).
5. Wilcoxon, R. *Some rapid approximate statistical procedures.* New York: American Cyanimid Co., 1949.

Journal of Experimental Psychology
Vol. 45, No. 5, 1953

EFFECTS OF TASK INSTRUCTIONS ON SOLUTION OF DIFFERENT CLASSES OF ANAGRAMS[1]

IRVING MALTZMAN AND LLOYD MORRISETT, JR.

University of California at Los Angeles

Previous experiments by the writers have explored implications of Hull's principle of habit strength and related corollaries for directed thinking (**5, 6**). The conditions under which habit strength plays a differential role in determining the direction of thinking are characteristic of those described as mental set. However, numerous experiments have shown that the course of thinking may be directed under conditions other than those producing the growth of habit strength (**4**). In particular, the instructions or task given *S* may have such an influence. According to the present formulation these conditions may give rise to a nonassociative factor contributing to the effective reaction potential determining the problem-solving behavior.

Rees and Israel (**7**) attempted to study the effects of task instructions on the solution of different five-letter anagrams. However, their study involved possible differences in training effects as well as the difference in instructions administered to experimental and control groups.

Although it does not appear possible to determine directly the effects of task instructions as distinguished from training in a problem-solving situation, this may be done indirectly. If the effects of task instructions plus training on set solutions are determined, and the effects of training alone are also observed, the approximate effects of instructions alone may be inferred. In an experimental study of this nature it would also be desirable to precede the set test trials by a training session so that increments in habit strength occurring during the test session would be relatively small.

The purpose of the present experiment was to determine the effects of task instructions upon the percentage of set solutions for different classes of anagrams by this indirect procedure, and to study the summation of the effects of training and task instructions.

METHOD

Subjects.—Fifty volunteers from introductory psychology classes served as *S*s in the experiment. There were ten *S*s in each of the five experimental conditions.

Procedure.—The training series for all conditions consisted of 20 five-letter anagrams having only one word as a solution which were typed in lower case on 3 × 5-in. index cards. A test series followed of 15 five-letter anagrams possessing several different words as solutions. Only one of the alternative solutions was a set solution related to the given class of training anagrams.

In the nature-training or N condition the solution of each of the anagrams in the training series was a word referring to some aspect of nature, for example, "fruit." The test anagrams had one solution referring to some aspect of nature and at least one other solution not directly or obviously related to nature. For example, the words "peach" or "cheap" would be solutions for the same anagram, but only the former would be a set solution. An attempt was made to increase the homogeneity of the nature class, and thereby avoid the development of competing sets (**6**), by employing only anagrams referring to plants and their habitat. The different anagrams in this condition were in no consistent letter order or arrangement.

A second condition was the order-training or O condition. The training anagrams in this condition had only one possible solution and all were in the same letter order, 32145. The third letter in the word was the first in the anagram, and so on. The anagrams were not consistently

[1] This investigation was supported in part by the Faculty Committee on Research, University of California, Southern Section.

related to each other in any other obvious manner. Anagrams in the test series all had multiple solutions one of which was in the 32145 letter order.

A third condition was the nature-instructions or NI condition. Anagrams in the training and test series were the same as those in the nature-training condition. This condition differed from the N condition only in the addition of special instructions. The fourth condition was the order-instructions or OI condition. It differed from the O condition only in the instructions administered. These four conditions formed a 2×2 factorial design.

A fifth condition was employed which was the compound of the first two conditions, order- and nature-training. Anagrams in the training and test series were the same as those in the N condition. In addition they were in the letter order of the O condition for the training series. In the test series the nature-word solution was in the 32145 letter order. The alternative solutions in the test series were in no consistent letter order.

Words used to form anagrams for the different conditions are shown in Table 1. Most of the words were obtained from lists used by Rees and Israel (**8**). A 2-min. time limit was enforced for each anagram in the training and test series of all conditions. During the training series, if *S* failed to obtain a correct solution in the allotted time, or gave an incorrect solution, *E* gave the correct solution. This correction technique was not employed during the test series. If *S* did not solve a test anagram within the allotted time or gave an incorrect solution, *E* presented another anagram without comment. Three practice anagrams were presented prior to the start of the experiment. They were in random letter orders and were not directly related to the class of nature anagrams. Instructions to the different groups were as follows: "This is an experiment on thinking. We want to find out how anagrams are solved. The results will in no way affect your class grades. You are to solve the five-letter anagrams that I will present to you as quickly as possible. A five-letter anagram is simply five scrambled letters. Your solution must use all the letters, that is, it must be a real five-letter word. It must not be a proper noun or slang. As soon as you have the answer, call it out to me and I will give you another to solve, but be sure your word fits the five letters before you name it. If you can't find a solution within two minutes, we will go on to another. Here is an example. What word can you make out of these five letters: c-i-k-b-r? The solution is brick. Now solve this one . . . and now this one. . . . Are there any questions? Do not ask any questions once the experiment has begun. Remember, call out your answer as soon as you find one."

After the presentation of the three practice anagrams, condition NI received the following additional task instructions: "During the experiment it will help you to look for words referring to flowers, growing things or their habitat. If you find some other solution you will also be credited with a correct answer."

Condition OI received the following task instructions after presentation of the practice anagrams: "During the experiment it will help you to look for a particular letter order. If you find some other solution you will also be credited with a correct answer."

TABLE 1

WORDS EMPLOYED AS ANAGRAMS

Practice: brick, chair, erase

Nature Training	*Order Training*
peony birch plant ferns	child cigar comes cream
lotus herbs fruit tulip	fresh cards waken house
lawns grape pansy hedge	towel mouth write uncle
grain prune phlox stems	dozen lends scarf lunch
bough daisy beans onion	hands learn party match
Nature Test	*Order Test*
aster, stares, tears, rates, tares	cause, sauce
range, anger	least, steal, stale, tales, slate
plums, slump, lumps	march, charm
peach, cheap	satin, stain, saint
pears, spear, pares, rapes, spare, parse, spear, reaps	baker, break, brake
	board, broad
	broth, throb
petal, leapt, pleat, plate	teach, cheat
thorn, north	trace, crate
stalk, talks	gates, stage
cedar, cared, raced	sword, words
palms, lamps	lamed, medal
pines, spine, penis, snipe	waits, waist
straw, warts, swart	wrote, tower
dates, stead	worth, throw
weeds, sewed	
trees, steer, terse, reset	

RESULTS AND DISCUSSION

The mean percentages of set solutions in the test series obtained under the first four experimental conditions are presented in Table 2. Bartlett's test of homogeneity gave a chi square of 1.45, $df = 3$, $p < .50$. The results of the analysis of variance of these four conditions are summarized in Table 3. The main effects of instructions and anagram class were signifi-

cant beyond the .05 and .01 levels of confidence, respectively. The latter finding is in accord with the results of a previous study of the writers (6). The interaction of anagram class and instructions was not significant. Thus the results indicate that task instructions may produce a significant increase in the percentage of set solutions independently of training effects.

The task instructions produced an increase of 6% in the nature condition and 12.7% in the order condition. These increases due to instructions are considerably less than the increments reported by Rees and Israel (7), indicating that the latter study may have confounded training and instruction effects.

The percentage of set solutions obtained in Cond. ON was 85.3, $s^2 = 225.6$. The significance of the difference between Cond. ON and N was tested by means of the t test, giving a t of 5.25, $df = 18$, $p < .01$. A comparison of Cond. ON and O yielded a t of 3.30, $df = 18$, $p < .01$. The latter t is greater than that obtained in a previous study (6).[2]

A comparison of Cond. ON with Cond. OI and NI would indicate the relative effects of the summation of two habit structures in Cond. ON with the summation of associative and nonassociative factors in Cond. OI and ON. The comparison of Cond. ON and OI gave a t of 1.39, $df = 18$, $p = .20$. The comparison of Cond. ON and NI gave a t of 3.97, $df = 18$, $p < .01$.

[2] The percentages of set solutions obtained in Cond. ON, O, and N were substituted in Hull's shrinkage index (2, 6). The index was 85.3/63.3 + 54.7 = 72.3. This value of the index is slightly higher than that previously obtained (6).

TABLE 2

MEAN PERCENTAGE OF SET SOLUTIONS FOR DIFFERENT CLASSES OF ANAGRAMS AND INSTRUCTIONS

Measure	Training		Task Instructions	
	Nature (N)	Order (O)	Nature (NI)	Order (OI)
Mean	54.7	63.3	60.7	76.0
S^2	106.9	219.7	162.3	219.2

TABLE 3

ANALYSIS OF VARIANCE OF SET SOLUTIONS

Source	df	Mean Square	F
Anagram class	1	1441.20	8.14**
Instructions	1	871.43	4.92*
Interaction	1	111.89	
Within groups	36	177.04	
Total	39		

* Significant at the .05 level.
** Significant at the .01 level.

According to Hull (2, 3) habit strength is a function of the number of reinforcements, and possibly other variables. Habit strength is the associative factor determining performance and has the characteristic of a state variable. It persists relatively unchanged with the passage of time. The present S–R formulation assumes that the phenomena described as mental set in directed thinking can be accounted for by the principle of habit strength and related corollaries in Hull's theory of behavior. However, the direction of thinking may be influenced by variables other than those producing habit strength. This study has demonstrated that task instructions may increase significantly the percentage of set solutions independently of the effects of training. This increase is assumed to be due to an increased amount of reaction potential produced by increments contributed by a nonassociative factor. In classical psychology this factor was the determining tendency or the Aufgabe (4). An analogous concept in Hull's theory of behavior is that of the fractional anticipatory goal response (1). On the basis of the present study, however, it is not possible to determine the characteristics of this nonassociative variable, or how it may be related to other intervening variables in Hull's theory. These are all problems of future research.

Summary

It was found that both instructions and anagram class significantly influenced the percentage of set solutions. The summation of order and nature classes of anagrams produced significantly more set solutions than either component class. The summation of order and nature produced significantly more set solutions than the summation of nature and task instructions but not order and task instructions.

The results of the experiment were discussed in terms of Hull's theory of behavior.

(Received September 10, 1952)

References

1. Hull, C. L. Goal attraction and directing ideas conceived as habit phenomena. *Psychol. Rev.*, 1931, **38**, 487–506.
2. Hull, C. L. *Principles of behavior*. New York: D. Appleton-Century, 1943.
3. Hull, C. L. *Essentials of behavior*. New Haven: Yale Univer. Press, 1951.
4. Humphrey, G. *Thinking*. New York: Wiley, 1951.
5. Maltzman, I., & Morrisett, L., Jr. Different strengths of set in the solution of anagrams. *J. exp. Psychol.*, 1952, **44**, 242–246.
6. Maltzman, I., & Morrisett, L., Jr. The effects of single and compound classes of anagrams on set solutions. *J. exp. Psychol.*, 1953, **45**, 345–350.
7. Rees, H., & Israel, H. An investigation of the establishment and operation of mental sets. *Psychol. Monogr.*, 1935, **46**, No. 6 (Whole No. 210).

Journal of Experimental Psychology
Vol. 45, No. 5, 1953

STUDIES OF DISTRIBUTED PRACTICE: XI. AN ATTEMPT TO RESOLVE CONFLICTING FACTS ON RETENTION OF SERIAL NONSENSE LISTS

BENTON J. UNDERWOOD[1]

Northwestern University[2]

In 1940 Hovland (3) published a report in which retention of serial nonsense lists was measured after 2 min., 10 min., and 24 hr. In one case lists were learned by massed practice (6-sec. intertrial interval) and in another by distributed practice (2-min. intertrial interval). The results clearly showed that retention as measured by both recall and relearning was better following distributed practice than following massed practice. We have reported a similar experiment (5) in which retention after 24 hr. was much better following massed practice than following distributed practice. Since the two studies are in direct conflict, it is important that attempts be made to determine the source of this conflict.

The most apparent difference between Hovland's experiment and ours was the difference in length of intertrial rests defining distributed practice. The longest intertrial rest we used was 1 min., as compared with the 2-min. interval used by Hovland. Consequently, in one of the experiments to be reported we have used 60, 90, and 120 sec. as intertrial rests. If length of rest interval is the critical difference between the two studies, we would expect sharp increments in retention between 60 and 120 sec.

The second experiment to be reported here extends the retention interval to 48 hr. In Hovland's experiment and in our previous experiment the maximum retention interval was 24 hr. It is conceivable that retention differences following massed and distributed practice increase or decrease as the retention interval increases and that these changes are related to length of intertrial interval defining distributed practice. In the second experiment, therefore, we have used intertrial intervals in learning of 2, 60, and 120 sec. with a 48-hr. retention interval. Both experiments may be compared roughly with our previous study in which the intertrial intervals were 2, 30, and 60 sec., and the retention interval 24 hr.

Procedure

General.—Each experiment consisted of three experimental conditions and one practice session. In Exp. I the only difference among conditions was length of intertrial rest during learning, these three intervals being 60, 90, and 120 sec. In Exp. II the intervals were 2, 60, and 120 sec. In Exp. I the retention interval was 24 hr.; in Exp. II, 48 hr. Original learning was always carried to one perfect recitation. Relearning was by massed practice regardless of interval used during learning and the criterion of relearning was one perfect trial.

Lists.—Each list consisted of 14 nonsense syllables having association values of 46.67% to 53.33% as calibrated by Glaze (2). Since the first syllable in a list was used only as an anticipatory cue, *S* actually learned only 13 items. The lists were constructed to have high intralist similarity. These lists are reproduced in a previous report (5) along with the rules used in constructing them. They were presented on Hull-type drums at a 2-sec. rate with learning taking place by the anticipation method.

[1] E. J. Archer supervised the gathering of the data; Jack Richardson is largely responsible for the statistical analysis.

[2] This work was done under Contract N7onr-45008, Project NR 154-057 between Northwestern University and the Office of Naval Research.

Specific conditions.—On the practice day *S* first learned the practice list to seven correct responses on a single trial by massed practice. Instructions were then given for the activity used to fill intertrial rests. This activity, symbol cancellation, has been described in detail elsewhere (**4**). Following these instructions, *S* continued learning to one perfect trial with 30 sec. between each trial. Five minutes following original learning the practice list was recalled and relearned. On the first experimental day *S* learned an experimental list to one perfect trial using the intertrial interval appropriate for that day. On the second day (Exp. I) the list was recalled and relearned. Immediately after relearning, the second experimental list was learned. This second experimental list was in turn recalled and relearned the following day, etc. The procedure was exactly the same for Exp. II except that 48 hr. elapsed between each session.

Each experiment had three conditions. Within each experiment both lists and intertrial rest conditions were completely counterbalanced against practice by using 36 *Ss*. The method for statistical analysis of such designs has been discussed by Archer (**1**). All *Ss* were undergraduate students.

Results

Practice day.—Since some direct comparisons between the two experiments will be made, it is necessary to demonstrate that performance of the two groups of *Ss* did not differ significantly on the practice list. The mean number of trials required to learn the practice list by *Ss* in Exp. I was 29.58, and for those in Exp. II, 28.22. The corresponding mean number of errors per trial was 1.38 and 1.41. For both trials and errors the within-group variance is greater than that between groups. The product-moment correlation between trials to learn the practice list and trials to learn all experimental lists combined for all 72 *Ss* was .58 ± .12. The value for errors was .66 ± .12.

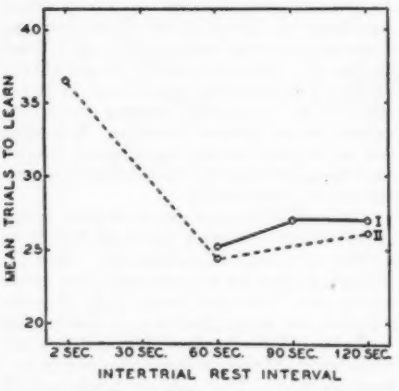

FIG. 1. Learning of serial nonsense lists as a function of intertrial interval

Learning.—The mean number of trials to learn under the different intertrial rest conditions is shown in Fig. 1. It may first be noted that in Exp. II distributed practice (as defined by 60- and 120-sec. rests) markedly facilitated learning as compared with massing (2-sec. rest). Secondly, it is observed that there is very little change in number of trials to learn between 60 and 120 sec. for either experiment. For Exp. I, *F* for trials to learn is less than one.

In the previous experiment (**5**) we used intertrial rest conditions of 2, 30, and 60 sec. and the same lists as used in the present experiments. Comparing the 2-sec. condition of the previous experiment and that of Exp. II shows the means are quite comparable. However, the mean trials to learn on the 60-sec. condition of the previous experiment (30.5) and on the 60-sec. conditions for the present two experiments (25.22 and 24.61) are quite diverse. The reason for this appears to lie in the selection of *Ss* which occurred in the present experiments. In these experiments the learning session was quite long even for an average *S* when the 2-min. distributed condition was used. A slow learner, requiring 40 trials to learn, would be in the experimental room about 1 hr. and 40 min. Consequently, we lost many *Ss* who were

TABLE 1

RETENTION OF SERIAL NONSENSE LISTS AS A FUNCTION OF INTERTRIAL INTERVAL DURING LEARNING*

Experiment	Intertrial Interval (Sec.)					σ_m	F
	2	30	60	90	120		
First Recall Trial							
I			4.28	5.28	4.72	.40	1.59
1952 (5)	7.72	4.08	4.36			.36	5.81
II	2.92		2.67		2.39	.33	.65
First and Second Recall Trials Combined							
I			9.94	12.06	11.03	.65	2.63
1952 (5)	13.64	9.94	10.28			.61	11.22
II	9.42		8.39		6.86	.60	4.60

* The scores are mean number of items recalled on first recall trial and on first and second recall trials combined. The estimate of σ_m is based on pooled *Ss* × practice variance. There are 2 and 66 *df* for each row; *F* at .05 level is 3.14; at .01, 4.95.

unable to complete the learning under the longer distribution intervals. The *Ss* we retained would be those most benefited by distributed practice, for these *Ss* could learn with a 2-min. distribution interval within a reasonable period of time. Therefore, we would expect the time to learn under the distributed conditions would be less in the present experiments than in the former.

Recall.—The mean number of items correct on the first recall trial and the mean number correct on the first and second recall trials combined are shown in Table 1. The entry, "1952 (5)," refers to the previously reported experiment (5) in which intervals of 2, 30, and 60 sec. were used with retention taken after 24 hr. Comparison of the results of Exp. I with those of the previous experiment cannot be made with absolute confidence since we do not know what the effect of subject selection (discussed above) may have on recall. It is apparent from Table 1, however, that recall following learning with 60-sec. intertrial interval (a condition which both experiments have in common) is very nearly the same on both the first recall trial and the first and second recall trials combined.

The essential facts given in Table 1 may be summarized as follows:

1. Experiment 1 shows no difference in recall for intervals of from 60 to 120 sec., and in all conditions recall is less than for the massed condition of the previous experiment. There is, then, no evidence that the difference between Hovland's findings and our previous findings is to be attributed to differences in length of distribution interval. Distribution intervals up to 2 min. give no evidence that distributed practice gives better recall than massed practice; indeed, all the evidence suggests that under the conditions of our research retention is better following massing than following distribution.

2. In Exp. II, where recall took place after 48 hr., the trend is likewise toward higher recall after massed practice than after distributed practice. On the first recall trial many

zero scores were recorded. However, the differences on the first and second recall trial combined almost attains a .01 level of confidence. Thus, there is no evidence that longer retention intervals and longer distribution intervals will change the basic generalization that in serial learning of nonsense syllables retention is better following massed practice than following distributed practice.

Relearning.—The mean number of trials to relearn for Exp. I was 11.86, 11.00, and 11.67 for 60-, 90-, and 120-sec. intertrial rests, respectively. The *F* is far from significant. In Exp. II, on the other hand, the values were 10.75, 11.83, and 13.72 for the 2-, 60-, and 120-sec. conditions. *F* is 4.00, which is approximately the .02 level of confidence. The most rapid relearning occurred with massed practice as would be expected from the recall scores.

Discussion

In conformance with two previous studies (**4, 5**) we have found that retention of serial nonsense lists is better following massed practice than following distributed practice. The results of the present studies make it unlikely that the discrepancy between Hovland's (**3**) findings (better recall following distributed practice) and ours is to be attributed to differences in intertrial intervals defining distribution or to length of retention interval. Therefore, the question still remains as to what could produce such markedly different results.

Our guess at the present time is that the contradiction is an indirect product of level of practice of *Ss* used in Hovland's research as compared with ours. Hovland's *Ss* served in many conditions and were given several practice days before they learned the 11-item experimental lists. The rate of learning the experimental lists was roughly twice that of our *Ss*. We have summarized elsewhere (**6**) the fragmentary evidence which suggests that items reinforced (correctly anticipated) many times will be recalled better following massed practice than following distributed practice, while items reinforced a few times will be better recalled following distributed practice. Well-practiced *Ss* (as in Hovland's research) correctly anticipate most items only a small number of times, whereas poorly practiced *Ss* (as in our research) have many items which have been reinforced many times. We might therefore expect Hovland's research to show better retention following distributed practice and ours better retention following massed practice. We hope to undertake research to determine if these guesses have any validity.

Summary

Hovland has found that following learning of serial nonsense lists by massed and distributed practice, retention is better for items learned under distribution. In more recent studies we found quite the opposite. One possible source of the discrepancy was the length of interval defining distributed practice, Hovland having used 2 min. as compared with 1 min. in our studies. Another possible factor was length of retention interval, it being conceivable that length of retention interval and length of distribution interval jointly determine whether retention will be facilitated by massed or by distributed practice. Two experiments were performed, one using distribution intervals of 60, 90, and 120 sec. between each learning trial with retention measured after 24 hr., and the other using intertrial intervals of 2, 60, and 120 sec., and a retention interval of 48 hr. Learning was car-

ried to one perfect trial under all conditions. The results show:

1. Better learning by distributed practice than by massed, but no change in rate of learning between 60- and 120-sec. intertrial intervals.
2. Better retention following massed practice than following distributed practice; this was evident for both recall and relearning. Little difference in retention was observed for intervals between 60 and 120 sec.
3. There was no indication that differences in retention following massed and distributed practice would decrease for retention intervals longer than 48 hr.

It is concluded that the differences between Hovland's results and our previous findings cannot be accounted for on the basis of length of intertrial interval defining distribution. It is possible that the differences between the studies may result from the fact that Hovland's *S*s were very highly practiced whereas ours were not.

(Received September 2, 1952)

References

1. Archer, E. J. Some greco-latin analysis of variance designs for learning studies. *Psychol. Bull.*, 1952, **49**, 521–537.
2. Glaze, J. A. The association value of nonsense syllables. *J. genet. Psychol.*, 1928, **35**, 255–269.
3. Hovland, C. I. Experimental studies in rote-learning theory: VI. Comparison of retention following learning to the same criterion by massed and distributed practice. *J. exp. Psychol.*, 1940, **26**, 568–587.
4. Underwood, B. J. Studies of distributed practice: VI. The influence of rest-interval activity in serial learning. *J. exp. Psychol.*, 1952, **43**, 329–340.
5. Underwood, B. J. Studies of distributed practice: VII. Learning and retention of serial nonsense lists as a function of intralist similarity. *J. exp. Psychol.*, 1952, **44**, 80–87.
6. Underwood, B. J. Studies of distributed practice: IX. Learning and retention of paired adjectives as a function of intralist similarity. *J. exp. Psychol.*, 1953, **45**, 197–203.

Journal of Experimental Psychology
Vol. 45, No. 5, 1953

THE MEMORY EFFECT OF VISUAL PERCEPTION OF THREE-DIMENSIONAL FORM

HANS WALLACH, D. N. O'CONNELL, AND ULRIC NEISSER

Swarthmore College

The kinetic depth effect that has been discussed in a preceding paper (2) enables monocular *Ss* to perceive three-dimensional form as directly as do persons with serviceable binocular vision by means of retinal disparity. Yet this effect alone does not, of course, solve the entire problem of the perception of solid form. Three-dimensional form is seen monocularly also when the observer does not move in relation to the object and it is also perceived in photographs and drawings. It has been mentioned in the preceding paper that an empiristic explanation of these cases of three-dimensional form perception becomes more feasible through the demonstration of the kinetic depth effect. This is so because no empiristic explanation can be termed successful until it is made clear how the original process or experience is brought about under whose influence current experience is supposed to occur. Prior to the demonstration of the kinetic depth effect (KDE) no process was known which could account in a satisfactory way for the "original" perception of three-dimensional form in monocular *Ss*.

Two different approaches have been made to explain the perception of three-dimensional form that occurs in the absence of retinal disparity or of other specific cues for visual depth. It has been proposed that three-dimensional forms are seen under these circumstances because the corresponding retinal patterns have the power to evoke them directly. Gibson (1), who holds such a view, believes that such retinal patterns have geometric characteristics which are specific stimuli for depth just as there are specific stimuli for color, pitch, etc. Many Gestalt psychologists believe that visual processes are spontaneously organized so that certain patterns of stimulation lead to three-dimensional forms and others to plane forms in perception and they have tried to formulate the principles which underlie such organization. When three-dimensional objects are seen as three-dimensional forms, it is due to the fact that their retinal projections have properties which favor organization as three-dimensional forms. The other approach is, of course, the empiristic one. It is believed that previous experiences can cause a present perception in three dimensions.

For a number of reasons, one of which—the nature of the KDE itself—will be discussed below, we came to believe that an influence of past experience plays an important role in the perception of three-dimensional form and set out to demonstrate such an effect in a stringent way. Such a demonstration requires that a retinal pattern, which at the outset is seen as a plane figure, gives rise under identical external conditions to the perception of a three-dimensional form after an intervening exposure of the same pattern given under conditions which cause it to be seen as three-dimensional.

METHOD

There are several ways in which a pattern can be made to appear as a three-dimensional form in the intervening exposure. We found it most appropriate to use the KDE for this purpose, and

experiments were done with a shadow technique (2). A three-dimensional wire figure was placed behind the translucent screen which was so chosen that its shadow visible on the other side of the screen looked two-dimensional to all *S*s in a stationary exposure. In the intervening exposure, the wire figure was turned back and forth so that a deforming shadow was cast which *S* eventually perceived as a three-dimensional form due to the operation of the KDE. This was followed sooner or later by the test exposure in which the same stationary shadow was presented which had been shown in the first exposure. The critical question of the experiment was whether or not it would now appear three-dimensional.

The same wire figures were employed as in the experiments of the preceding paper (2). They were the "helix," the "parallelogram," and the "110° corner." Three experimental series will be reported and in each all three figures were presented. However, the purpose of experimenting with the 110° corner was different from the purpose of experimentation with the other two figures. Merely for technical reasons were all three figures presented in the same series. The results for the 110° corner will therefore be discussed separately at the end of this article. We begin with a description of the three series. Since each of the series, insofar as the helix and the parallelogram were concerned, was designed to answer a different question, the results of the different series will be separately presented and discussed.

The experimental series were composed of moving exposures and of stationary presentations of the various figures shown individually. In the moving exposure the wire figure behind the screen rotated back and forth through an angle of 42° at the rate of one cycle in 1.5 sec. The shadow was shown for 10 sec. by turning the lamp in whose light the shadow was visible on and off, and then *S* was asked for a report on what he had seen. The 10-sec. exposure was repeated until the report clearly indicated that *S* had seen the correct three-dimensional form. If this did not happen within 12 such exposures, the moving presentation was listed as having failed.

The shadow for the stationary presentation was cast by the wire figure in a position within the range of the rotation of 42°. Figures 2, 3A, and 4 of the previous article (2) show the shadows employed in the stationary presentations. All stationary presentations employed the same shadow of each wire figure and lasted for 5 sec. Our *S*s were undergraduates of Swarthmore College.

Series I.—There was first a stationary exposure of the 110° corner. Then, moving exposures of the helix, the 110° corner, and the parallelogram were given. Stationary test exposures of the three figures in the same order followed. Thus, the moving exposure and the test exposure of a given figure occurred a few minutes apart. For the helix, the moving exposure of the 110° corner and of the parallelogram intervened, and, in the case of the parallelogram, the stationary exposures of the helix and of the 110° corner came between the moving and the test exposure of that figure. This series was presented to 33 *S*s.

Since no stationary exposure of the helix and the parallelogram was made in this series prior to the moving exposure, stationary presentations of these figures were given to a control group of 16 *S*s to find out what percentage of *S*s would see the figures two-dimensional at the outset.

Series II.—Here two groups of 20 *S*s were employed. Only the 110° corner and *one* of the other two figures were presented in moving exposure. The remaining figure served as a control and was given only in the final stationary exposure together with the two experimental figures. Thus, for an *S* of the helix group the sequence of presentation was the following: First came a stationary exposure of the helix followed by one of the 110° corner. Thereupon, the moving presentations of the helix and of the 110° corner were made in that order. These were followed by stationary exposures of the helix, of the 110° corner, and finally of the parallelogram which had not been seen by these *S*s before. For *S*s of the other group, the parallelogram was given in place of the helix and vice versa.

Series III.—This series differed from Series I only in two points. The sequence in which the figures were presented both in the moving and in the stationary test exposure was parallelogram, helix, and 110° corner. Also, a large time interval was introduced between the moving and the test exposures. For 12 *S*s this interval was 24 hr. and for 11 *S*s seven days.

Results for "Helix" and "Parallelogram"

The purpose of Series I was simply to demonstrate that the perception of a three-dimensional form in the moving exposure would tend to make the figure appear three-dimensional in the stationary test exposure also. Evidence that prior to a moving exposure our two figures appear two-dimensional in stationary presentation comes here from a control group of 16 *S*s. Both figures were seen plane by all of them.

The results for the moving and for the test exposure are simple in the case of the parallelogram. All 33 *Ss* saw this figure three-dimensional in the moving exposure and all of them reported the same three-dimensional form in the stationary test exposure. In the case of the helix, 31 of the 33 *Ss* reported this figure as three-dimensional in the moving presentation after various exposure times; the remaining 2 saw it as a plane deforming figure and saw it two-dimensional in the test exposure also. Of the critical 31 *Ss* who had seen the helix three-dimensional in the moving presentation a majority of 26 reported seeing the three-dimensional form when the stationary shadow of the figure was presented in the test exposure; the remaining 5 *Ss* saw a plane figure.

These results demonstrate a strong influence of an earlier experience in the perception of three-dimensional form. What is the nature of this influence? Does it consist in a tendency to see further figures three-dimensional after some have been perceived in this fashion under the same circumstances, or is it an influence of an earlier perception of a particular figure on the perceptual process which takes place when this particular figure is given again? Series II was designed to answer this question. It will be remembered that only one of the two figures was presented to a given *S* in moving exposure, but both figures were presented in the test exposure. If we are dealing with a general tendency, both figures should be seen three-dimensional in the test exposure by a majority of the *Ss*. If the influence is in the nature of an individual figure causing a later exposure of that figure to be seen in the same three-dimensional way, only the figure previously given in moving exposure should appear three-dimensional.

In this series a stationary control exposure of the figure for which an aftereffect was to be established was given to the experimental *Ss* prior to the moving presentations. All 40 *Ss* saw the figure that was presented to them as being two-dimensional and this confirmed the results of 16 control *Ss* of Series I.

In the moving exposure, 18 out of 20 *Ss* saw the helix three-dimensional and all *Ss* perceived the parallelogram in this fashion. In the stationary test presentation, 13 out of the 18 *Ss* who had been given a moving exposure of the helix and had then seen a three-dimensional form saw this figure again three-dimensional, and 5 reported a plane figure. The parallelogram was seen three-dimensional by 17 of the 20 "parallelogram *Ss*." These data agree well with those obtained in Series I and show a strong influence of the perception of three-dimensional form in the moving presentation on the reports in the test exposure. The new information to be gained from the present series comes from the test exposures of the figures which had *not* been previously presented to the respective *S*. Of the 20 *Ss* who had been given the parallelogram in moving exposure and had seen it three-dimensionally, only one perceived the helix in three dimensions. The results are somewhat different for the parallelogram in this situation. As many as 7 of the 20 *Ss* to whom it had not been presented in moving exposure reported seeing it in three dimensions (19 of them had seen at least the 110° corner three-dimensional).

For the helix, these results are quite unequivocal. Whereas 13 out of 18 *Ss* who had seen this figure three-dimensional before saw it so in the text exposure, this was the case with only 1 out of the 20 *Ss* who had seen only the other figures in this fashion.

The influence which causes the perception of three-dimensional form in the stationary test exposure appears to come from a previous three-dimensional perception of the *same* figure only. For the parallelogram we have to compare the result 17 out of 20 *Ss* who had seen the same figure three-dimensional with 7 out of 19 *Ss* who had seen only either one or two other figures in this fashion. Although there is for this figure an influence of a general kind, the specific influence which comes from a previous three-dimensional perception of the same figure is much stronger. It should be mentioned that the shadow of the parallelogram is more easily seen in three dimensions than that of the helix. Often a mere suggestion like "could this be a tetrahedron" suffices to make this figure appear three-dimensional. When the results for the helix and the parallelogram are taken together, the difference between the cases due to a general effect (8 out of 40) and those cases where the same figure had been seen three-dimensional before (30 out of 38) is reliable at better than the .01 level of confidence. It seems safe to conclude that our aftereffect consists in an influence of the perception of a figure upon a subsequent perceptual process which takes place when the same figure is given again, and that, in the case of some figures, previous exposures of different figures may exert a similar influence as if by suggestion. (Further evidence on this point comes from the experiments with the 110° corner reported below.)

Some readers may find difficulty with this formulation. How do we know, they may ask, that we have here really an effect on perception, that *Ss* actually *saw* a three-dimensional form? It could be that *Ss* reported a three-dimensional form because they had previously seen the same pattern on the screen as three-dimensional and knew that the pattern represented such a form. The evidence on this point is clear and simple. When directly after the test exposure *Ss* were allowed to inspect the stationary shadow for a longer period, a large percentage of them reported Necker cube-like reversals of the figures that they had described. Everybody who has seen a drawing of a three-dimensional figure reverse will agree that only an actually perceived three-dimensional form will exhibit these changes. Where these reversals occur, coming on unexpectedly and initially appearing to be objective, three-dimensional forms are seen with all concreteness. Many *Ss* of Series II who had reported seeing three-dimensional forms on the test presentation were given prolonged exposures of the stationary figures. For the helix, 13 out of 14, and for the parallelogram, 20 out of 24 *Ss* reported reversals spontaneously.

In Series I and II, the time interval between the perception of a figure as a three-dimensional form and the test presentation amounted to a few minutes. In Series III this time interval was much longer. Twelve *Ss* were tested after 24 hr. and 11 after seven days. All *Ss* perceived a three-dimensional form in the moving exposure of both figures. In the 24-hr. group, all *Ss* saw the parallelogram three-dimensionally in the test exposure and 11 out of 12 did so in case of the helix. In the seven-day group these numbers are 10 out of 11 for both figures.

In this series the test for the parallelogram always preceded that of the helix, and this may have favored a three-dimensional appearance of the helix. We therefore report data from another similar experiment where, after a 24-hr. interval, the helix was the first figure tested. Of nine *Ss* who had seen the helix three-dimensional in the moving exposure, seven saw it three-dimensional in the test presentation.

It appears from these data that the aftereffect can be obtained virtually undiminished after longer time intervals, and that it should be termed a memory effect. Moreover, the results of Series II indicated that we were dealing largely with the effect of individual memory traces.

How does a memory trace produce its effect

on the perception in a test exposure? Does it merely give an indication that the pattern concerned must be perceived as three-dimensional or has a trace the capacity to determine a specific three-dimensional form for the new perception? So far, the answer to this question comes only from a consideration of the kinetic depth effect (KDE) which was described in the previous paper (2). The very nature of this process makes it necessary to ascribe to a memory trace the power to determine the organization of a visual form process. Stimulation for the KDE consists in a deforming retinal projection which is produced when a three-dimensional object changes its orientation to *S*. At any moment the retinal projection assumes a slightly different shape and every one of these momentary images can have a form such that it would produce a perception of a two-dimensional figure, if it were presented by itself, that is, not in the context of the deforming projection. However, within the context, that is, when it is given in continuous sequence with all the other momentary images which make up the deforming projection, it produces the perception of a three-dimensional object which changes its apparent orientation to *S*. The apparent momentary aspect of the perceived object corresponds to the particular retinal image which is given at that moment. This makes it clear that any one of the momentary retinal images gives rise to a perceived three-dimensional form only because it was preceded by a number of different images of the object. At the moment when it is given on the retina the preceding images are matters of the past. Pertinent stimulation is given in temporal sequence and the perceived form is its cumulative result. From the moment at which three-dimensional form is first perceived, a complex memory trace which represents this result of the preceding stimulation must be assumed to participate in the perceptual process. To be sure, this trace alone is not the correlate of perceptual experience; stimulation by one of the momentary retinal images is also necessary to bring about the three-dimensional percept. But it is obvious that the perceptual process to which a momentary image gives rise must attain its form in three-dimensional space due to such a trace.

If perceived three-dimensional form must be ascribed to the organizing power of a trace in the case of the KDE, it seems justifiable to assume that the effect of a trace in our test exposure is of the same nature. The motion of any one of our wire figures in kinetic presentation can be stopped and *S* will continue to see the stationary shadow as the three-dimensional form that he had seen during the moving presentation. There appears to be no reason to assume that at this point, a trace action of a different nature takes place. One should rather think that here, too, the trace causes perception of a specific three-dimensional form. It may be mentioned that such consideration of the nature of the KDE was one of the reasons why we expected to find a wide influence of past experience on form perception.

Results for "110° Corner"

We have shown that a memory effect on form perception is easily demonstrated and that it readily bridges large time intervals. We are inclined to believe that this effect plays a large role in ordinary perception of space and of solid form. We think that, at least in the adult, memory effects are responsible for the majority of instances of perception of solid form and of the spatial arrangement of the objects in the visual field.

If this is so, why did we demonstrate the effect only with two different figures? The answer is that suitable wire figures are hard to find because the shadows of most of them look three-dimensional from the outset. We believe that this is so because of the great wealth of previous experience with three-dimensional form. A pattern, like Fig. 1, that can be interpreted as representing three surfaces meeting to form a spatial corner, will always be seen in that fashion. This is true even in cases where such a

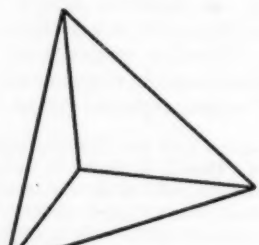

Fig. 1. Figure always seen as three-dimensional

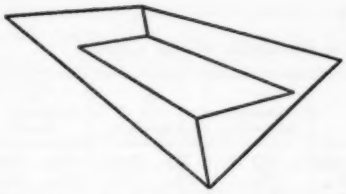

FIG. 2. Figure that makes no sense as a three-dimensional form

pattern is part of a large one as in Fig. 2. Although this figure as a whole makes no sense as a three-dimensional form to most *Ss*, three-dimensionality is seen at the two places where three lines meet to form a **Y**, most convincingly so at the lower one. Such a **Y**-shaped pattern is very frequently present in the projections of solid objects, as, for instance, in the projections of corners of boxes and of rooms, and there is much occasion to establish previous experience of three-dimensionality in connection with it. Wire figures which contain it are not suitable for our experiments where the given shadow pattern ought to appear two-dimensional in the control exposure.

Yet, there is nevertheless a way to make use of such a pattern for the demonstration of a memory effect, namely, if we are satisfied with establishing a modified three-dimensional form in the moving exposure. A **Y** pattern will look like the edges of a corner to many *Ss*, but that corner will frequently appear rectangular, presumably because in the great majority of previous encounters the **Y** pattern has been produced by a rectangular corner and has been seen as one. If, in the moving exposure, the **Y** pattern can be made to appear, say, as an obtuse corner, there is then the question of how it will look in the test exposure. If it is again seen as an obtuse corner, an aftereffect of previous experience has been established within our experiment.

Such an experiment was performed with the 110° corner. As mentioned above, this figure was presented in all three experimental series and in every one the control exposure was given to the experimental *Ss*. Thus, all *Ss* were given first a stationary control exposure, then a moving presentation, and finally a stationary test exposure of this figure. Fortunately, this particular **Y** pattern was seen as a two-dimensional figure in the control exposure by as many as 56 of the 96 *Ss* who participated in the three series, a number much higher than we had hoped for when the experiment had been planned; their results will be reported later. Of the remaining 40 *Ss*, 17 saw an obtuse corner in the control exposure and 23 saw a rectangular one. With regard to the question just raised, we are concerned with this latter group. In the case of 2 *Ss* of this group, the moving exposure failed to produce a perception of an obtuse corner, which reduces the number of *Ss* in this pertinent group to 21. Of these 21 *Ss* who did see a rectangular corner in the control exposure and an obtuse corner in the moving exposure, 8 reported a rectangular corner in the test presentation and 13 an obtuse corner. While the former did not display an aftereffect of perception in the moving exposure, the 13 *Ss* who saw an obtuse corner in the test presentation saw the shadow as one three-dimensional form at the outset and later, presumably as a memory effect of the moving exposure, as a different one. It means that a single experience can modify a form perception which itself may well be the outcome of previous experiences.

The total result of the experiment with the 110° corner is given in Table 1. It lists the results of all three

TABLE 1

NUMBER OF *Ss* GIVING VARIOUS RESPONSE SEQUENCES TO 110° CORNER

	A No Change			B Full Aftereffect			C Partial Aftereffect	D No Aftereffect		E Random Result	
Control	2D	Rect.	Obt.	2D	2D	Rect.	2D	2D	Rect.	Obt.	Obt.
Moving	2D	Rect.	Obt.	Rect.	Obt.	Obt.	Obt.	Obt.	Obt.	Obt.	Obt.
Test	2D	Rect.	Obt.	Rect.	Obt.	Obt.	Rect.	2D	Rect.	2D	Rect.
Series I	0	0	5	1	8	4	7	1	3	2	2
Series II	1	1	4	1	19	4	3	5	1	0	1
Series III	0	1	2	0	9	5	0	1	4	0	1
Total	1	2	11	2	36	13	10	7	8	2	4
Grand Total	14			51			10	15		6	

experimental series and is arranged in the following manner: Above the center line there are listed the 11 occurring combinations of *Ss*' reports in the three exposures. They play the role of headings for the 11 columns of data below and they are grouped under 5 capital letters in a manner which will be explained later on. The first horizontal row refers to the first stationary or control exposure; the second to the moving exposure; and the third row to the second stationary or test exposure. For instance, an *S* who sees the corner as a two-dimensional figure in the control exposure, as an obtuse corner in the moving exposure, and as a rectangular corner in the test exposure falls under the combination heading below the letter C(2D, Obt., Rect.), and the number of *Ss* who gave this particular sequence of reports in each series is given under this heading. The first three horizontal rows of numbers give the results for the three series separately and the fourth row the totals.

All three combination headings under A show no change from one exposure to another and the 14 *Ss* who gave these report sequences did not contribute in any way to the outcome of the experiment. This reduces the total number of *Ss* to be considered to 82.

The headings under B have two features in common: the reports for the moving exposure differ from those for the control exposures in the direction of a more adequate perception of the spatial form of the corner, and the reports for the moving and the test exposure are the same. Thus, these three headings represent three report sequences which denote a memory effect of the experience in the moving exposure on perception in the test exposure. However, to the 51 *Ss* who gave one of these report sequences should be added the 10 *Ss* under C who, although they did not give the same report in the test exposure that they gave in the moving exposure, nevertheless showed a change in the proper direction, if the control and the test reports are compared; they reported a plane figure in the former and a rectangular corner in the latter.

The two report sequences under D represent those cases that could have shown a memory effect of the perception in the moving exposure but did not; that is, the reports in the test exposure were the same as the reports in the control exposure. Under E two report sequences are listed which

make no sense, and the total of six *Ss* listed here gives an idea to what degree randomness figures in our results.

It should be pointed out that the data in the third horizontal row (Series III) were obtained with intervals of 24 hr. or seven days between the moving and the test exposures. The results for this group of 23 *Ss* do not differ significantly from those obtained after brief intervals (Series I and II).

To summarize: Of 82 *Ss*, 61 gave reports indicating a memory effect operating between perception in the moving and perception in the test exposure, 15 showed no such effect, and 6 gave results which denote independence of the perception in the test exposure of both preceding exposures. With three-quarters of the *Ss* yielding positive results, this experiment represents another demonstration of a memory effect on form perception.

The data given in Table 1 also throw light on an issue that has already been raised with regard to the two other wire figures, namely whether the aftereffect is caused by a previous three-dimensional perception of the *same* figure, or whether the figure is seen three-dimensional because of a general tendency arising in the experimental situation to see further figures three-dimensional. If the latter were true, one might expect that those *Ss* who see the corner three-dimensional in the test presentation as the result of an aftereffect will report a rectangular or an obtuse corner in about the same proportion as do *Ss* who see this figure three-dimensional at the outset. Among the 96 *Ss*, 40 saw the corner three-dimensional in the control exposure and of these 23 reported a rectangular and 17 an obtuse corner. On the other hand, a total of 48 *Ss* saw this figure three-dimensional in the test presentation as an aftereffect of the moving exposure; they are listed under the following three headings: 2D, Rect., Rect.; 2D, Obt., Obt.; and 2D, Obt. Rect. If the aftereffect consists in a general tendency to see the figure three-dimensional, one should expect these *Ss* to report a rectangular or an obtuse corner in about the same ratio as the previously mentioned 40 *Ss*, that is, in a proportion of about 23 to 17. This was not the case; only 12 *Ss* reported a rectangular corner and 36 an obtuse corner. The difference between 12 and 36 is reliably different from the difference between 23 and 17 at the .01 level of confidence.

If, on the other hand, the appearance of the corner in the test presentation largely depends on the perception of the corner in the moving exposure, we should expect to find a preponderance of *Ss* reporting an obtuse corner in the test, because 46 of the 48 *Ss* saw an obtuse corner in the moving exposure. This was indeed the case; 36 of the 46 *Ss* reported an obtuse corner again in the test presentation and the hypothesis that these two results are related is confirmed at the .01 level of confidence. In other words, these results can be ascribed, at least in part, to an influence of a specific perceptual experience upon a later perceptual process.

Summary

The shadows of three different three-dimensional wire figures were shown on a translucent screen. These figures were so chosen that their shadows appeared two-dimensional to the majority of *Ss*. By use of the kinetic depth effect, that is, by turning the wire figures back and forth, the shadows were then made to appear three-dimensional. After intervals which ranged from minutes to a week, the

stationary shadows were presented again in the same fashion in which they had been exposed originally and were then reported to appear three-dimensional by a large number of *Ss*. In order to make certain that these reports were based on three-dimensional percepts, rather than on inferences concerning the perceptual objects, a number of *Ss* were given prolonged test exposures and nearly all of them reported spontaneously reversals of the kind which are usually demonstrated with a Necker cube. Thus it was demonstrated that a previous perceptual experience can cause a later form perception to be three-dimensional. Evidence was presented that general set played only a minor role in the reported experiments and that the aftereffects obtained were to a large part due to the influence of the memory of individual figures.

(Received August 11, 1952)

References

1. Gibson, J. J. *The perception of the visual world.* Boston: Houghton Mifflin, 1950.
2. Wallach, H., & O'Connell, D. N. The kinetic depth effect. *J. exp. Psychol.*, 1953, **45**, 205–217.

Encyclopedia of

ABERRATIONS

A PSYCHIATRIC HANDBOOK

Edited by **EDWARD PODOLSKY, M.D.**

State University of New York Medical College

With a Foreword by **ALEXANDRA ADLER, M.D.**

New York University College of Medicine

This is the first systematic exposition of human aberrational behavior. In this volume over fifty eminent psychologists and psychiatrists discuss all types of aberrations, with particular emphasis on their psychodynamics. The material is arranged in alphabetical sequence for easy reference.

SOME OF THE ENTRIES:

Abasia
Ablutomania
Abulia
Acalculia
Acataphasia
Aggression
Alcoholism
Amnesia
Anal eroticism
Anancasm
Anti-Semitic attitudes
Anxiety, dental
Aphasia and linguistics
Autism, infantile
Auto-punishment
Benzedrine, addiction
Bestiality
Body image disturbances
Boredom
Brontophobia
Cacodaemonomania
Chloral delirium
Choreomania
Clairvoyance
Claustrophobia
Cocaine, addiction
Crime, neurotic
Criminality
Depression
Devil worship
Dream murders
Dysprosody
Ecouteur
Ecstasy, artificial
Erotographomania
Exhibitionism
Family tension
Fellatio
Fetishism
Folie à deux
Frigidity
Frottage
Gambling
Gammacism
Gelasmus
Gustatory sweating
Gynephobia
Hair-plucking
Hallucinations
Haptodysphoria
Hashish, addiction
Head banging
Heroin, addiction
Heterolalia
Homosexuality
Hysteria
Iconolagny
Illusions
Inferiority feelings
Intellectual malfunctioning
Kainotophobia
Kakorrhaphiophobia
Kleptomania
Language frustration
Laughter, fits of
Lesbianism
Logorrhea
Lying
Malingering
Masochism
Menstrual anomalies
Mescaline intoxication
Murderer, mind of
Mutism
Mysophobia
Narcolepsy
Necrophilia
Negativism
Nudism
Nymphomania
Ochlophobia
Onanism
Opium, addiction
Pavor nocturnus
Pessimism
Pethidine, addiction
Phobias
Pornography
Psychosis
Puberty, aberrational
Sadism
Schizophrenia
Somnambulism
Sophomania
Suicide
Therioanthropy
Xenophobia
Zoophilism

OVER HALF A MILLION WORDS

$10.00

PHILOSOPHICAL LIBRARY, *Publishers*

15 East 40th St., Desk 188 **New York 16, N. Y.**

Expedite shipment by prepayment

PSYCHOLOGICAL MONOGRAPH

#324
1951

A recent psychological monograph of general interest

VISUAL PERCEPTION AND THE ROTATING TRAPEZOIDAL WINDOW

BY

ADELBERT AMES, JR.

$1.00

Order from

American Psychological Association
1333 Sixteenth Street N.W.
Washington 6, D. C.